Problem set 4

Darragh

2023-04-16

Problem set 4

Question 1 We're interested in modeling the historical causes of child mortality. We have data from 26855 children born in Skellefte a, Sweden from 1850 to 1884. Using the "child" dataset in the eha library, fit a Cox Proportional Hazard model using mother's age and infant's gender as covariates. Present and interpret the output.

```
# remove objects
rm(list=ls())
# detach all libraries
detachAllPackages <- function() {
  basic.packages <- c("package:stats", "package:graphics", "package:grDevices", "package:utils", "packag
  package.list <- search()[ifelse(unlist(gregexpr("package:", search()))==1, TRUE, FALSE)]
  package.list <- setdiff(package.list, basic.packages)
  if (length(package.list)>0)  for (package in package.list) detach(package,  character.only=TRUE)
}
detachAllPackages()

# load libraries
pkgTest <- function(pkg){
  new.pkg <- pkg[!(pkg %in% installed.packages()[,  "Package"])]
  if (length(new.pkg))
    install.packages(new.pkg,  dependencies = TRUE)
  sapply(pkg,  require,  character.only = TRUE)
}

# here is where you load any necessary packages
# ex: stringr
 lapply(c("stringr"),  pkgTest)
```

```
## Loading required package: stringr
```

```
## [[1]]
## stringr
##    TRUE
```

```
lapply(c("survival", "eha", "tidyverse", "ggfortify", "stargazer"),  pkgTest)
```

```
## Loading required package: survival
```

```
## Loading required package: eha
```

```
## Loading required package: tidyverse
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0     v purrr   1.0.1
## v tibble  3.1.8     v dplyr   1.1.0
## v tidyr   1.3.0     v forcats 1.0.0
## v readr   2.1.3
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## Loading required package: ggfortify
##
## Loading required package: stargazer
##
##
## Please cite as:
##
##
##  Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.
##
##  R package version 5.2.3. https://CRAN.R-project.org/package=stargazer
```

```
## [[1]]
## survival
##     TRUE
##
## [[2]]
##  eha
## TRUE
##
## [[3]]
## tidyverse
##      TRUE
##
## [[4]]
## ggfortify
##      TRUE
##
## [[5]]
## stargazer
##      TRUE
```

```
## Load library
library(ggplot2)
library(tidyverse)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
##
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
library(eha)
```

```
data("child")
dat <- child
```

Overview of the data

```r
glimpse(dat)
```

```
## Rows: 26,574
## Columns: 10
## $ id        <int> 9, 150, 158, 178, 263, 342, 363, 393, 408, 486, 497, 563, 57~
## $ m.id      <int> 246606, 377744, 118277, 715337, 978617, 282943, 341341, 8408~
## $ sex       <fct> male, male, male, male, female, male, male, male, female, fe~
## $ socBranch <fct> farming, farming, worker, farming, worker, farming, farming,~
## $ birthdate <date> 1853-05-23, 1853-07-19, 1861-11-17, 1872-11-16, 1855-07-19,~
## $ enter     <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
## $ exit      <dbl> 15.000, 15.000, 15.000, 15.000, 0.559, 0.315, 15.000, 15.000~
## $ event     <dbl> 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, ~
## $ illeg     <fct> no, no, no, no, no, no, no, no, no, yes, no, no, no, no, no,~
## $ m.age     <dbl> 35.009, 30.609, 29.320, 41.183, 42.138, 32.931, 42.360, 28.6~
```

```r
child_surv <- with(dat, Surv(enter, exit, event))
km <- survfit(child_surv ~ 1, data = dat)
summary(km, times = seq(0, 15, 1))
```

```
## Call: survfit(formula = child_surv ~ 1, data = dat)
##
##  time n.risk n.event censored survival std.err lower 95% CI upper 95% CI
##     0  26574       0        0    1.000 0.00000        1.000        1.000
##     1  24319    2161       94    0.919 0.00168        0.915        0.922
##     2  23450     778       91    0.889 0.00193        0.885        0.893
##     3  22766     596       88    0.867 0.00209        0.862        0.871
##     4  22269     430       68    0.850 0.00220        0.846        0.854
##     5  21859     365       44    0.836 0.00228        0.832        0.841
##     6  21533     261       65    0.826 0.00233        0.822        0.831
##     7  21266     214       53    0.818 0.00238        0.813        0.823
##     8  21077     151       38    0.812 0.00241        0.807        0.817
##     9  20915     117       45    0.808 0.00243        0.803        0.812
##    10  20777     103       35    0.804 0.00245        0.799        0.808
##    11  20655      81       41    0.801 0.00246        0.796        0.805
##    12  20531      91       33    0.797 0.00248        0.792        0.802
##    13  20404      89       38    0.794 0.00250        0.789        0.798
##    14  20277      95       32    0.790 0.00251        0.785        0.795
##    15  20141      84    20193    0.787 0.00253        0.782        0.792
```

```r
plot(km, main = "Kaplan-Meier Plot", xlab = "Years", ylim = c(0.7, 1))
```

ps4_files/figure-latex/unnamed-chunk-3-1.pdf

```r
autoplot(km)
```

ps4_files/figure-latex/unnamed-chunk-3-2.pdf

```
### effcts
##mage not working as continous, so bins
dat$m.age <-cut(dat$m.age, breaks=c(15, 20,25,30, 35, 40,45, 50), right = FALSE)
kn_mage <- survfit(child_surv ~ m.age, data = dat)
autoplot(kn_mage)
```

ps4_files/figure-latex/unnamed-chunk-4-1.pdf

```
#####mage
cox_mage <- coxph(child_surv ~ m.age , data = dat)
summary(cox_mage)
```

```
## Call:
## coxph(formula = child_surv ~ m.age, data = dat)
##
##   n= 26572, number of events= 5616
##    (2 observations deleted due to missingness)
##
##                    coef exp(coef) se(coef)      z Pr(>|z|)
## m.age[20,25) -0.35771   0.69928  0.11380 -3.143  0.00167 **
## m.age[25,30) -0.27175   0.76204  0.11047 -2.460  0.01389 *
## m.age[30,35) -0.29498   0.74455  0.11039 -2.672  0.00754 **
## m.age[35,40) -0.22375   0.79952  0.11094 -2.017  0.04372 *
## m.age[40,45) -0.15254   0.85852  0.11396 -1.339  0.18072
## m.age[45,50) -0.04396   0.95700  0.15206 -0.289  0.77253
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## m.age[20,25)    0.6993      1.430    0.5595    0.8740
## m.age[25,30)    0.7620      1.312    0.6137    0.9463
## m.age[30,35)    0.7445      1.343    0.5997    0.9244
## m.age[35,40)    0.7995      1.251    0.6433    0.9937
## m.age[40,45)    0.8585      1.165    0.6867    1.0734
## m.age[45,50)    0.9570      1.045    0.7104    1.2893
##
## Concordance= 0.518  (se = 0.004 )
## Likelihood ratio test= 26.95  on 6 df,   p=1e-04
## Wald test            = 27.67  on 6 df,   p=1e-04
## Score (logrank) test = 27.76  on 6 df,   p=1e-04
```

```
drop1(cox_mage, test = "Chisq")
```

```
## Single term deletions
##
## Model:
## child_surv ~ m.age
##        Df    AIC    LRT  Pr(>Chi)
## <none>    113014
## m.age   6 113029 27.911 9.765e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
stargazer(cox_mage, type = "text")
```

```
## 
## ================================================
##                          Dependent variable:
##                      ---------------------------
##                              child_surv
## ------------------------------------------------
## m.age[20,25)                  -0.358***
##                                (0.114)
## 
## m.age[25,30)                  -0.272**
##                                (0.110)
## 
## m.age[30,35)                  -0.295***
##                                (0.110)
## 
## m.age[35,40)                  -0.224**
##                                (0.111)
## 
## m.age[40,45)                   -0.153
##                                (0.114)
## 
## m.age[45,50)                   -0.044
##                                (0.152)
## 
## ------------------------------------------------
## Observations                   26,572
## R2                              0.001
## Max. Possible R2                0.986
## Log Likelihood               -56,500.790
## Wald Test                27.670*** (df = 6)
## LR Test                  26.951*** (df = 6)
## Score (Logrank) Test     27.764*** (df = 6)
## ================================================
## Note:                *p<0.1; **p<0.05; ***p<0.01
```

```r
exp(-0.358)###20,25
```

```
## [1] 0.6990731
```

```r
exp(-0.272)###25,30
```

```
## [1] 0.7618543
```

```r
exp(-0.295)###30,35
```

```
## [1] 0.7445316
```

```r
exp(-0.224)###35,40
```

```
## [1] 0.7993151
```

The coefficients in the output show the hazard ratio for each age category relative to the reference category (15- 20), after adjusting for infant gender. For example, the hazard ratio for ages between 20 and 25 is exp(-0.358) = 0.698, which means that the hazard (risk) of death for infants born to mothers in this age group is 0.698 times the hazard for infants born to mothers under age 20, after adjusting for gender. Similarly, the hazard ratio for ages between 25 and 30 is 0.762, indicating a lower risk of death compared to the reference group, and the hazard ratio for ages between 30 and 35 is 0.744, also indicating a lower risk of death.

The standard errors in parentheses indicate the precision of the estimates. The R-squared value is very low (0.001), indicating that the model explains very little of the variability in the data.

Overall, this output suggests that infants born to mothers in older age categories have a lower risk of mortality than infants born to younger mothers, after adjusting for gender. However, the effect sizes are small, and other factors not included in the model may have a stronger influence on child mortality

```
##sex graph
kn_s <- survfit(child_surv ~ sex, data = dat)
autoplot(kn_s)
```

ps4_files/figure-latex/unnamed-chunk-5-1.pdf

```
####sex
cox_s <- coxph(child_surv ~ sex , data = dat)
summary(cox_s)
```

```
## Call:
## coxph(formula = child_surv ~ sex, data = dat)
##
##   n= 26574, number of events= 5616
##
##                coef exp(coef) se(coef)      z Pr(>|z|)
## sexfemale -0.08332   0.92005  0.02674 -3.116  0.00183 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## sexfemale    0.9201      1.087    0.8731    0.9696
##
## Concordance= 0.511  (se = 0.003 )
## Likelihood ratio test= 9.72  on 1 df,   p=0.002
## Wald test            = 9.71  on 1 df,   p=0.002
## Score (logrank) test = 9.71  on 1 df,   p=0.002
```

```
drop1(cox_s, test = "Chisq")
```

```
## Single term deletions
##
## Model:
## child_surv ~ sex
##        Df    AIC    LRT Pr(>Chi)
## <none>    113022
## sex     1 113029 9.7232  0.00182 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
stargazer(cox_s, type = "text")
```

```
##
## ================================================
##                          Dependent variable:
##                      ---------------------------
##                              child_surv
## ------------------------------------------------
## sexfemale                     -0.083***
##                                (0.027)
##
## ------------------------------------------------
## Observations                   26,574
## R2                             0.0004
## Max. Possible R2                0.986
## Log Likelihood               -56,509.880
## Wald Test                 9.710*** (df = 1)
## LR Test                   9.723*** (df = 1)
## Score (Logrank) Test      9.715*** (df = 1)
## ================================================
## Note:                *p<0.1; **p<0.05; ***p<0.01
```

```
exp(-0.083)## female
```

```
## [1] 0.9203511
```

The coefficient in the output shows the hazard ratio for females relative to males, after adjusting for mother's age. In this case, the hazard ratio for females is exp(-0.083) = 0.920, which means that the hazard (risk) of death for female infants is 0.920 times the hazard for male infants, after adjusting for mother's age.

The standard error in parentheses indicates the precision of the estimate. The R-squared value is very low (0.0004), indicating that the model explains very little of the variability in the data.

Overall, this output suggests that male infants have a slightly higher risk of mortality than female infants, after adjusting for mother's age. However, the effect size is small, and other factors not included in the model may have a stronger influence on child mortality.

```
##sex and mage additive
#########could be improved by seperating data by gender
kn_s_mage_a <- survfit(child_surv ~ sex + m.age, data = dat)
autoplot(kn_s_mage_a)
```

ps4_files/figure-latex/unnamed-chunk-6-1.pdf

```
##sex and mage interactive
cox_s_mage_i <- coxph(child_surv ~ sex * m.age, data = dat)
summary(cox_s_mage_i)
```

```
## Call:
## coxph(formula = child_surv ~ sex * m.age, data = dat)
##
##   n= 26572, number of events= 5616
##    (2 observations deleted due to missingness)
##
##                              coef exp(coef) se(coef)      z Pr(>|z|)  
## sexfemale                 -0.08493   0.91858  0.21558 -0.394   0.6936  
## m.age[20,25)              -0.34799   0.70611  0.15361 -2.265   0.0235 *
## m.age[25,30)              -0.26398   0.76799  0.14892 -1.773   0.0763 .
## m.age[30,35)              -0.28169   0.75451  0.14872 -1.894   0.0582 .
## m.age[35,40)              -0.26492   0.76727  0.14963 -1.770   0.0766 .
## m.age[40,45)              -0.12734   0.88043  0.15351 -0.830   0.4068  
## m.age[45,50)              -0.05899   0.94271  0.20997 -0.281   0.7787  
## sexfemale:m.age[20,25) -0.01534   0.98478  0.22872 -0.067   0.9465  
## sexfemale:m.age[25,30) -0.01163   0.98844  0.22208 -0.052   0.9583  
## sexfemale:m.age[30,35) -0.02486   0.97545  0.22193 -0.112   0.9108  
## sexfemale:m.age[35,40)  0.08955   1.09369  0.22302  0.402   0.6880  
## sexfemale:m.age[40,45) -0.05140   0.94990  0.22912 -0.224   0.8225  
## sexfemale:m.age[45,50)  0.03763   1.03835  0.30494  0.123   0.9018  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                        exp(coef) exp(-coef) lower .95 upper .95
## sexfemale                 0.9186     1.0886    0.6020    1.4016
## m.age[20,25)              0.7061     1.4162    0.5225    0.9542
## m.age[25,30)              0.7680     1.3021    0.5736    1.0283
## m.age[30,35)              0.7545     1.3254    0.5637    1.0099
## m.age[35,40)              0.7673     1.3033    0.5722    1.0288
## m.age[40,45)              0.8804     1.1358    0.6517    1.1895
## m.age[45,50)              0.9427     1.0608    0.6247    1.4227
## sexfemale:m.age[20,25)    0.9848     1.0155    0.6290    1.5418
## sexfemale:m.age[25,30)    0.9884     1.0117    0.6396    1.5275
## sexfemale:m.age[30,35)    0.9754     1.0252    0.6314    1.5070
## sexfemale:m.age[35,40)    1.0937     0.9143    0.7064    1.6933
## sexfemale:m.age[40,45)    0.9499     1.0527    0.6062    1.4883
## sexfemale:m.age[45,50)    1.0383     0.9631    0.5712    1.8876
##
## Concordance= 0.522  (se = 0.004 )
## Likelihood ratio test= 39.61  on 13 df,   p=2e-04
## Wald test            = 40.28  on 13 df,   p=1e-04
## Score (logrank) test = 40.45  on 13 df,   p=1e-04
```

```
drop1(cox_s_mage_i, test = "Chisq")
```

```
## Single term deletions
##
## Model:
## child_surv ~ sex * m.age
##           Df    AIC    LRT Pr(>Chi)
## <none>       113015                
## sex:m.age  6 113006 3.2212   0.7806
```

```
stargazer(cox_s_mage_i, type = "text")
```

```
##
## ==================================================
##                              Dependent variable:
##                          ---------------------------
##                                 child_surv
## --------------------------------------------------
## sexfemale                         -0.085
##                                   (0.216)
##
## m.age[20,25)                     -0.348**
##                                   (0.154)
##
## m.age[25,30)                      -0.264*
##                                   (0.149)
##
## m.age[30,35)                      -0.282*
##                                   (0.149)
##
## m.age[35,40)                      -0.265*
##                                   (0.150)
##
## m.age[40,45)                      -0.127
##                                   (0.154)
##
## m.age[45,50)                      -0.059
##                                   (0.210)
##
## sexfemale:m.age[20,25)            -0.015
##                                   (0.229)
##
## sexfemale:m.age[25,30)            -0.012
##                                   (0.222)
##
## sexfemale:m.age[30,35)            -0.025
##                                   (0.222)
##
## sexfemale:m.age[35,40)             0.090
##                                   (0.223)
##
## sexfemale:m.age[40,45)            -0.051
##                                   (0.229)
##
## sexfemale:m.age[45,50)             0.038
##                                   (0.305)
##
## --------------------------------------------------
## Observations                      26,572
## R2                                 0.001
## Max. Possible R2                   0.986
## Log Likelihood                  -56,494.460
## Wald Test                     40.280*** (df = 13)
## LR Test                       39.610*** (df = 13)
## Score (Logrank) Test          40.447*** (df = 13)
## ==================================================
## Note:                  *p<0.1; **p<0.05; ***p<0.01
```

```
# exponentiate parameter estimates to obtain hazard ratios
exp(-0.348)###20,25  and male
```

```
## [1] 0.7060989
```

```
exp(-0.264)###25,30 and male
```

```
## [1] 0.7679735
```

```
exp(-0.282)###30,35 and male
```

```
## [1] 0.7542737
```

```
exp(-0.265)###35,40 and male
```

```
## [1] 0.7672059
```

```
##additive
cox_s_mage_a <- coxph(child_surv ~ sex + m.age, data = dat)
summary(cox_s_mage_a)
```

```
## Call:
## coxph(formula = child_surv ~ sex + m.age, data = dat)
##
##   n= 26572, number of events= 5616
##    (2 observations deleted due to missingness)
##
##                  coef exp(coef) se(coef)      z Pr(>|z|)
## sexfemale    -0.08211   0.92117  0.02675 -3.070  0.00214 **
## m.age[20,25) -0.35535   0.70093  0.11380 -3.123  0.00179 **
## m.age[25,30) -0.26956   0.76371  0.11047 -2.440  0.01468 *
## m.age[30,35) -0.29332   0.74579  0.11039 -2.657  0.00788 **
## m.age[35,40) -0.22327   0.79990  0.11094 -2.012  0.04417 *
## m.age[40,45) -0.15104   0.85981  0.11396 -1.325  0.18504
## m.age[45,50) -0.04048   0.96033  0.15206 -0.266  0.79011
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## sexfemale       0.9212      1.086    0.8741    0.9708
## m.age[20,25)    0.7009      1.427    0.5608    0.8761
## m.age[25,30)    0.7637      1.309    0.6150    0.9483
## m.age[30,35)    0.7458      1.341    0.6007    0.9259
## m.age[35,40)    0.7999      1.250    0.6436    0.9942
## m.age[40,45)    0.8598      1.163    0.6877    1.0750
## m.age[45,50)    0.9603      1.041    0.7128    1.2938
##
## Concordance= 0.522  (se = 0.004 )
## Likelihood ratio test= 36.39  on 7 df,   p=6e-06
## Wald test            = 37.1  on 7 df,   p=4e-06
## Score (logrank) test = 37.19  on 7 df,   p=4e-06
```

```
drop1(cox_s_mage_a, test = "Chisq")
```

```
## Single term deletions
##
## Model:
## child_surv ~ sex + m.age
##        Df    AIC    LRT  Pr(>Chi)
## <none>    113006
## sex     1 113014  9.4379 0.0021254 **
## m.age   6 113022 27.6259 0.0001105 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
stargazer(cox_s_mage_a, type = "text")
```

```
## 
## ================================================
##                           Dependent variable:   
##                       ---------------------------
##                               child_surv        
## ------------------------------------------------
## sexfemale                     -0.082***         
##                                (0.027)          
##                                                 
## m.age[20,25)                  -0.355***         
##                                (0.114)          
##                                                 
## m.age[25,30)                  -0.270**          
##                                (0.110)          
##                                                 
## m.age[30,35)                  -0.293***         
##                                (0.110)          
##                                                 
## m.age[35,40)                  -0.223**          
##                                (0.111)          
##                                                 
## m.age[40,45)                   -0.151           
##                                (0.114)          
##                                                 
## m.age[45,50)                   -0.040           
##                                (0.152)          
##                                                 
## ------------------------------------------------
## Observations                   26,572           
## R2                              0.001           
## Max. Possible R2                0.986           
## Log Likelihood               -56,496.070        
## Wald Test                37.100*** (df = 7)     
## LR Test                  36.389*** (df = 7)     
## Score (Logrank) Test     37.194*** (df = 7)     
## ================================================
## Note:                *p<0.1; **p<0.05; ***p<0.01
```

```
exp(-0.082)##female
```

```
## [1] 0.921272
```

```
exp(-0.355)##20-25
```

```
## [1] 0.7011734
```

```
exp(-0.270)#25,30
```

```
## [1] 0.7633795
```

```
exp(-0.293)#30,35
```

```
## [1] 0.7460221
```

```
exp(-0.223)#35,40
```

```
## [1] 0.8001148
```

Based on the results of the Cox proportional hazards models you ran, we can make the following conclusions:

Age is significantly associated with child survival. Children in the age group 20-25 have a significantly higher hazard of child death compared to the reference age group (0-20). Similarly, children in the age groups 25-30, 30-35, and 35-40 have significantly higher hazards of child death compared to the reference group. Sex is also significantly associated with child survival. Female children have a significantly lower hazard of child death compared to male children. There is evidence of an interaction between sex and age on child survival. However, the coefficients for the interaction terms are not significant, indicating that the effect of age on child survival does not differ significantly between male and female children. When looking at the hazard ratios for the age groups, we see that the hazard of child death decreases with increasing age, meaning that older children are less likely to die than younger children. When looking at the hazard ratio for sex, we see that female children have a 8% lower hazard of child death than male children.

The Cox proportional hazards model is a type of survival analysis model used to examine the association between predictor variables and time-to-event outcomes. In this case, the outcome of interest is child survival, and the predictor variables include the mother's age group and sex.

The parameter estimates obtained from the Cox proportional hazards models for the two models are as follows:

Model with interaction term: The hazard ratio for females compared to males is exp(-0.085) = 0.919. The hazard ratio for the age group 20-25 compared to the reference group (under 20) for males is exp(-0.348) = 0.706. The hazard ratio for the age group 25-30 compared to the reference group (under 20) for males is exp(-0.264) = 0.768. The hazard ratio for the age group 30-35 compared to the reference group (under 20) for males is exp(-0.282) = 0.754. The hazard ratio for the age group 35-40 compared to the reference group (under 20) for males is exp(-0.265) = 0.767. Model without interaction term: The hazard ratio for females compared to males is exp(-0.082) = 0.921. The hazard ratio for the age group 20-25 compared to the reference group (under 20) for males is exp(-0.355) = 0.701. The hazard ratio for the age group 25-30 compared to the reference group (under 20) for males is exp(-0.270) = 0.763. The hazard ratio for the age group 30-35 compared to the reference group (under 20) for males is exp(-0.293) = 0.746. The hazard ratio for the age group 35-40 compared to the reference group (under 20) for males is exp(-0.223) = 0.800. In general, a hazard ratio less than 1 indicates a lower risk of the event (in this case, child mortality) for the group being compared to the reference group, while a hazard ratio greater than 1 indicates a higher risk.

Comparing the two models, we can see that the estimates for the age groups are similar, but the estimates for the effect of sex differ slightly. In the model without an interaction term, the hazard ratio for females is slightly higher (0.921) compared to the model with the interaction term (0.919), indicating a slightly higher risk of child mortality for females. However, the difference is small and may not be practically significant.