Statistics 641, Fall 2011
Homework #1
Answers

1. Given the two-by-two table:

Treatment	Dead	Alive	total
A	6	19	25
B	16	11	27
	22	30	52

let ψ be the odds ratio for association between treatment and mortality. The null hypothesis is $H_0\colon \psi = 1$.

(a) Compute the Pearson chi-square statistic for H_0.

The expected values under H_0 are

Treatment	Dead	Alive	total
A	$22 \times 25/52 = 10.58$	$30 \times 25/52 = 11.42$	25
B	$22 \times 27/52 = 11.42$	$22 \times 27/52 = 15.58$	27
	22	30	52

$$\frac{(6 - 25 \times 22/52)^2}{22 \times 30 \times 25 \times 27/52^3} = 6.62$$

(b) Compute the maximum likelihood estimate of $\beta = \log \psi$ and its variance.

let ψ be the odds ratio for association between

$$\hat{\beta} = \log \frac{16 \times 19}{6 \times 11} = 1.527$$

$$\text{var}(\hat{\beta}) = \frac{1}{16} + \frac{1}{19} + \frac{1}{6} + \frac{1}{11} = 0.3727$$

(c) Find a 95% confidence interval for ψ.

A 95% CI for $\log \psi$ is

$$1.527 \pm \sqrt{0.3727} \times 1.96 = (0.331, 2.724)$$

so a 95% CI for ψ is

$$(e^{0.331}, e^{2.724}) = (1.392, 15.240)$$

(d) Compute the Wald test-statistic for H_0.

$$\frac{\hat{\beta}^2}{\text{var}(\hat{\beta})} = \frac{1.527^2}{0.3727} = 6.259$$

2. Do Problem 7.2 from the textbook. Please perform the calculations by hand and show your work.

(a)

Letting R_j be the number at risk at each time, d_j the number of events, and $\hat{S}(t_j)$ estimate of the survivor function. For groups A and B we have:

A

t_j	R_j	d_j	$1 - d_j/R_j$	$\hat{S}(t_j)$
0	10	–	–	1.00
23	6	1	5/6	0.833
24	5	1	4/5	0.667
26	4	1	3/4	0.500
28	3	1	2/3	0.333
30	2	1	1/2	0.167
31	1	1	0/1	0.00

B

t_j	R_j	d_j	$1 - d_j/R_j$	$\hat{S}(t_j)$
0	10	–	–	1.00
9	10	1	9/10	0.900
12	9	1	8/9	0.800
13	8	1	7/8	0.700
14	7	2	5/7	0.500
16	5	1	4/5	0.400

Combined

t_j	R_j	d_j	$1 - d_j/R_j$	$\hat{S}(t_j)$
0	20	–	–	1.00
9	19	1	18/19	0.947
12	17	1	16/17	0.892
13	16	1	15/16	0.836
14	15	2	13/15	0.724
16	13	1	12/13	0.669
23	8	1	7/8	0.585
24	6	1	5/6	0.488
26	5	1	4/5	0.390
28	4	1	3/4	0.293
30	2	1	1/2	0.146
31	1	1	0/1	0.000

(b)

Combining both treatment groups:

	t_k	d_{k1}	n_{k1}	d_{k2}	n_{k2}	$n_{k1}+n_{k2}$ $(=w_k)$	$E[d_{k1}]$	$\mathrm{Var}(d_{k1})$
	9	0	9	1	10	19	$1\times 9/19$	$1\times 9\times 10\times 18/19^2\times 18$
	12	0	8	1	9	17	$1\times 8/17$	$1\times 8\times 9\times 16/17^2\times 16$
	13	0	8	1	8	16	$1\times 8/16$	$1\times 8\times 8\times 15/16^2\times 15$
	14	0	8	2	7	14	$2\times 8/15$	$2\times 8\times 7\times 13/15^2\times 14$
	16	0	8	1	5	13	$1\times 8/13$	$1\times 8\times 5\times 12/13^2\times 12$
	23	1	6	0	2	8	$1\times 6/8$	$1\times 6\times 2\times 7/8^2\times 7$
	24	1	5	0	1	6	$1\times 5/6$	$1\times 5\times 1\times 5/6^2\times 5$
	26	1	4	0	1	5	$1\times 4/5$	$1\times 4\times 1\times 4/5^2\times 4$
	28	1	3	0	1	4	$1\times 3/4$	$1\times 3\times 1\times 3/4^2\times 3$
	30	1	2	0	0	2	$1\times 2/2$	$1\times 2\times 0\times 1/2^2\times 1$
	31	1	1	0	0	1	$1\times 1/1$	$1\times 1\times 0\times 0/1^2\times 0$
$\sum(\cdot)$		6					8.26	2.12
$\sum w_k(\cdot)$		26					70	394

so we have for the unweighted log-rank:

$$\frac{(6-8.26)^2}{2.12} = 2.41$$

and for the Gehan-Wilcoxon-weighted log-rank:

$$\frac{(26-70)^2}{394} = 4.91$$

3. Using the data from problem 2, assume that the failure times follow an exponential distribution and

   (a) compute the score test for difference between groups.

The score test statistic is

$$(d_A - \hat{\lambda}T_A)^2\left(\frac{1}{\hat{\lambda}T_0}+\frac{1}{\hat{\lambda}T_A}\right) \sim \chi^2_1$$

d_A is the number of failures in group A, $\hat{\lambda}$ is the common hazard rate, T_j is the total follow-up time in group j. We have $d_A = 6$, $T_A = 215$, $d_B = 6$, $T_A = 171$, so

$$\hat{\lambda} = \frac{6+6}{215+171} = 0.0311$$

so

$$(d_A - \hat{\lambda}T_A)^2\left(\frac{1}{\hat{\lambda}T_A}+\frac{1}{\hat{\lambda}T_B}\right) = (6-0.0311\times 215)^2(\frac{1}{.0311\times 215}+\frac{1}{.0311\times 171}) = 0.158$$

(b) compute the hazard ratio and a 95% confidence interval.

The hazard ratio is simply

$$\frac{6/171}{6/215} = 1.257$$

The variance is best computed on the log-hazard-ratio scale, and is the inverse of the Fisher information,

$$\frac{1}{\hat{\lambda}_A T_A} + \frac{1}{\hat{\lambda}_B T_B} = \frac{1}{d_A} + \frac{1}{d_B} = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

On the log-scale, the CI is therefore $\log(1.257) \pm \sqrt{1/3}Z_{.975} = \log(1.257) \pm 1.132$, so on the HR-scale, we have $1.257e^{\pm 1.132} = (0.406, 3.898)$.

The variance can also be found using the delta-method.

Again, please perform the calculations by hand and show your work.

4. For the following use the data from the file `hw1.csv` (available on-line at `http://www.biostat.wisc.edu/~cook/641.homework.html`). These data can be read into `R` using a command such as

```
hw1 <- read.csv("hw1.csv")
```

(or use a dataset name of your choosing).

These data can be read into `SAS` using the command

```
proc import datafile="hw1.csv" dbms=csv out=hw1 ;
```

The variables in the dataset are:

`trt`	Treatment group (0/1)
`days`	Follow-up time in days
`status`	censoring/failure indicator (1=failure, 0=censored)
`sex`	Sex (1=Male, 2=Female)
`age`	Age at baseline in years

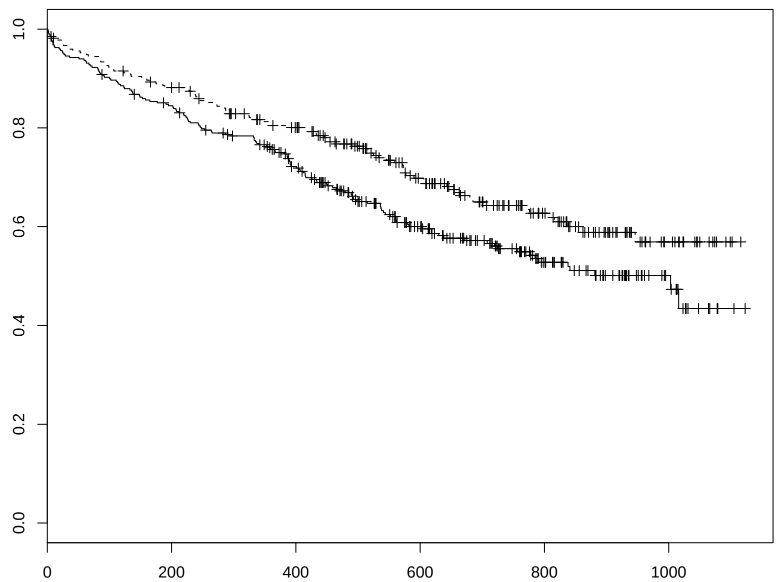
Let H_0 be the null hypothesis that there is no difference in survival by treatment.

First, read data into `R`:

```
> D <- read.csv("hw1.csv")
```

(a) Plot the Kaplan-Meier estimates of event-free survival by treatment group.

```
> plot(survfit(Surv(days,status)~trt, data=D), lty=1:2)
```



(b) Assess whether the failure times follow an exponential distribution.

This is most easily done by fitting a Weibull model, and considering the scale parameter.

```
> summary(survreg(Surv(days,status)~trt, data=D))

Call:
survreg(formula = Surv(days, status) ~ trt, data = D)
             Value Std. Error     z       p
(Intercept) 7.256     0.1108 65.50 0.00000
trt         0.376     0.1601  2.35 0.01892
Log(scale)  0.176     0.0598  2.95 0.00319

Scale= 1.19
...
```

The p-value for the test of log-scale $= 0$, is small (0.003) providing strong evidence that the data do not follow an exponential distribution. Note that the scale parameter is greater than 1, suggesting that the underlying hazard function is decreasing with time.

(c) Test H_0 using the Wald, score and likelihood ratio tests.

Fitting a Cox proportional hazards model,

```
> summary(coxph(Surv(days,status)~trt, data=D))
Call:
coxph(formula = Surv(days, status) ~ trt, data = D)

  n= 622

       coef exp(coef) se(coef)     z Pr(>|z|)
trt -0.3231    0.7239   0.1335 -2.42   0.0155 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

    exp(coef) exp(-coef) lower .95 upper .95
trt    0.7239      1.381    0.5573    0.9404

Rsquare= 0.01   (max possible= 0.99 )
Likelihood ratio test= 5.98  on 1 df,   p=0.01445
Wald test            = 5.86  on 1 df,   p=0.01552
Score (logrank) test = 5.91  on 1 df,   p=0.01507
```

so the Wald statistic is $Z = -2.42$ (or the chi-square statistic is $5.86 = Z^2$, with 1 DF), and the likelihood ratio statistic is 5.98 (chi-square with 1 DF). The score test is the log-rank which yields a statistic of 5.91. All are close together and provide modest evidence of a difference between treatment groups (and are quite similar to the results from the Weibull model).

(d) Provide a summary and 95% confidence interval for the observed treatment difference.

The hazard ratio estimate from the Cox-model above is HR = 0.7239 (0.5573, 0.9404)

(e) Test H_0 adjusted for age and sex using the Wald and likelihood ratio tests.

First fit the Cox-model including age and sex effects:

```
> summary(coxph(Surv(days,status)~trt + age + sex, data=D))
Call:
coxph(formula = Surv(days, status) ~ trt + age + sex, data = D)

  n= 622

         coef exp(coef)  se(coef)      z Pr(>|z|)
trt -0.337584  0.713492  0.133585 -2.527  0.01150 *
age  0.021732  1.021970  0.007538  2.883  0.00394 **
sex -0.082668  0.920657  0.173959 -0.475  0.63463
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

    exp(coef) exp(-coef) lower .95 upper .95
trt    0.7135     1.4016    0.5491     0.927
age    1.0220     0.9785    1.0070     1.037
sex    0.9207     1.0862    0.6547     1.295

Rsquare= 0.023   (max possible= 0.99 )
Likelihood ratio test= 14.54  on 3 df,   p=0.002257
Wald test            = 14.23  on 3 df,   p=0.002613
Score (logrank) test = 14.31  on 3 df,   p=0.002512
```

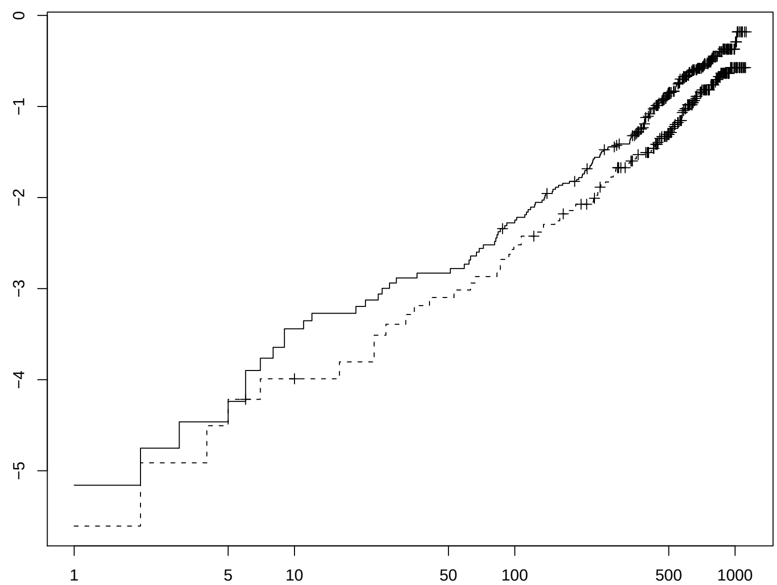
The Z statistic from the Wald test is $Z = -2.527$. To derive the likelihood ratio test we can simply take the difference between the likelihood ratio statistics for the models with and without treatment, or we can use the anova function:

```
> anova(coxph(Surv(days,status)~ age + sex, data=D),
+    coxph(Surv(days,status)~trt + age + sex, data=D), test="Chis")
Analysis of Deviance Table
 Cox model: response is  Surv(days, status)
 Model 1: ~ age + sex
 Model 2: ~ trt + age + sex
   loglik  Chisq Df P(>|Chi|)
1 -1426.1
2 -1422.8 6.5277  1   0.01062 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The adjusted likelihood-ratio statistic is 6.5277.

(f) Assess whether the proportional hazards assumption for the treatment difference is reasonable.

```
> plot(survfit(Surv(days,status)~trt, data=D), lty=1:2, fun="cloglog")
```



These curves remain roughly the same distance apart for the portion where they are most stable—there is no evidence from the plot that the PH assumption does not hold. Using the `cox.zph` function for the unadjusted and adjusted models, we have

```
> cox.zph(coxph(Surv(days,status)~ trt, data=D))
       rho chisq     p
trt 0.0279 0.185 0.667
> cox.zph(coxph(Surv(days,status)~ trt + age + sex, data=D))
          rho   chisq     p
trt    0.0296 0.20868 0.648
age    0.0024 0.00132 0.971
sex    0.0154 0.05721 0.811
GLOBAL     NA 0.27726 0.964
```

Neither suggest that there is evidence the PH assumption fails.

In SAS, the PH test can be run as follows:

```
proc import datafile="hw1.csv" dbms=csv out=hw1 ;

proc phreg  data = hw1;
   model days*status(0) = trt  daystrt ;
   daystrt = trt*log(days);
```

with selected output:

```
             Analysis of Maximum Likelihood Estimates

              Parameter   Standard                            Hazard
Variable DF    Estimate      Error Chi-Square Pr > ChiSq     Ratio

trt       1    -0.43342    0.51810     0.6998     0.4028     0.648
daystrt   1     0.02098    0.09502     0.0488     0.8252     1.021
```

The Wald chi-square statistic is 0.0488 (p=.82), so, again, there is no evidence that the PH assumption fails. (Note that the estimate for treatment in the above is not meaningful because of the presence of the time-treatment interaction term.)