Medical Statistics – Answers lab 9

Part 1: Analysis of overall survival in the Worcester Heart Attack study

Association between MI order and overall survival

Kaplan-Meier survival curves and logrank test

```
Call: survfit(formula = Surv(lenfol, fstat_numeric) ~ miord, data = whas500)

                miord=first
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
    1    329       5    0.985 0.00674       0.9717        0.998
    2    324       7    0.964 0.01034       0.9435        0.984
    3    317       1    0.960 0.01074       0.9397        0.982
    4    316       2    0.954 0.01150       0.9321        0.977
    5    314       1    0.951 0.01186       0.9284        0.975
    6    313       4    0.939 0.01317       0.9137        0.965
    7    309       3    0.930 0.01406       0.9029        0.958
   10    306       1    0.927 0.01434       0.8994        0.956
   11    305       2    0.921 0.01487       0.8923        0.951
   14    303       2    0.915 0.01538       0.8852        0.946
   16    301       1    0.912 0.01563       0.8817        0.943
   17    300       1    0.909 0.01587       0.8782        0.940
   18    299       2    0.903 0.01634       0.8713        0.935
   19    297       2    0.897 0.01678       0.8644        0.930
   22    295       1    0.894 0.01700       0.8609        0.928
   33    294       2    0.888 0.01742       0.8540        0.922
   34    292       1    0.884 0.01762       0.8506        0.920
   37    291       1    0.881 0.01782       0.8472        0.917
   42    290       1    0.878 0.01802       0.8438        0.914
   46    289       1    0.875 0.01821       0.8404        0.912
```

57	288	1	0.872	0.01840	0.8370	0.909
61	287	1	0.869	0.01858	0.8336	0.906
64	286	1	0.866	0.01877	0.8303	0.904
69	285	1	0.863	0.01894	0.8269	0.901
81	284	1	0.860	0.01912	0.8235	0.898
83	283	1	0.857	0.01929	0.8202	0.896
88	282	1	0.854	0.01946	0.8168	0.893
93	281	1	0.851	0.01963	0.8134	0.890
95	280	1	0.848	0.01979	0.8101	0.888
97	279	1	0.845	0.01995	0.8068	0.885
100	278	1	0.842	0.02011	0.8034	0.882
108	277	1	0.839	0.02027	0.8001	0.880
109	276	1	0.836	0.02042	0.7968	0.877
113	275	1	0.833	0.02057	0.7935	0.874
116	274	1	0.830	0.02072	0.7902	0.871
117	273	1	0.827	0.02087	0.7868	0.869
134	272	1	0.824	0.02101	0.7835	0.866
135	271	1	0.821	0.02115	0.7802	0.863
137	270	1	0.818	0.02129	0.7770	0.860
140	269	1	0.815	0.02143	0.7737	0.858
145	268	1	0.812	0.02156	0.7704	0.855
146	267	1	0.809	0.02169	0.7671	0.852
169	266	2	0.802	0.02195	0.7605	0.847
187	264	1	0.799	0.02208	0.7573	0.844
192	263	1	0.796	0.02220	0.7540	0.841
200	262	1	0.793	0.02232	0.7507	0.838
233	261	1	0.790	0.02244	0.7475	0.836
235	260	1	0.787	0.02256	0.7442	0.833
259	259	2	0.781	0.02279	0.7377	0.827
269	257	1	0.778	0.02291	0.7345	0.824
274	256	1	0.775	0.02302	0.7312	0.822
287	255	1	0.772	0.02313	0.7280	0.819
297	254	1	0.769	0.02324	0.7248	0.816
313	253	1	0.766	0.02334	0.7215	0.813
343	252	1	0.763	0.02345	0.7183	0.810
345	251	1	0.760	0.02355	0.7151	0.807
358	250	1	0.757	0.02365	0.7119	0.805
359	249	2	0.751	0.02385	0.7054	0.799
363	247	1	0.748	0.02394	0.7022	0.796
382	241	1	0.745	0.02405	0.6989	0.793
392	237	1	0.741	0.02415	0.6956	0.790
397	236	1	0.738	0.02425	0.6923	0.787
405	231	1	0.735	0.02435	0.6889	0.784

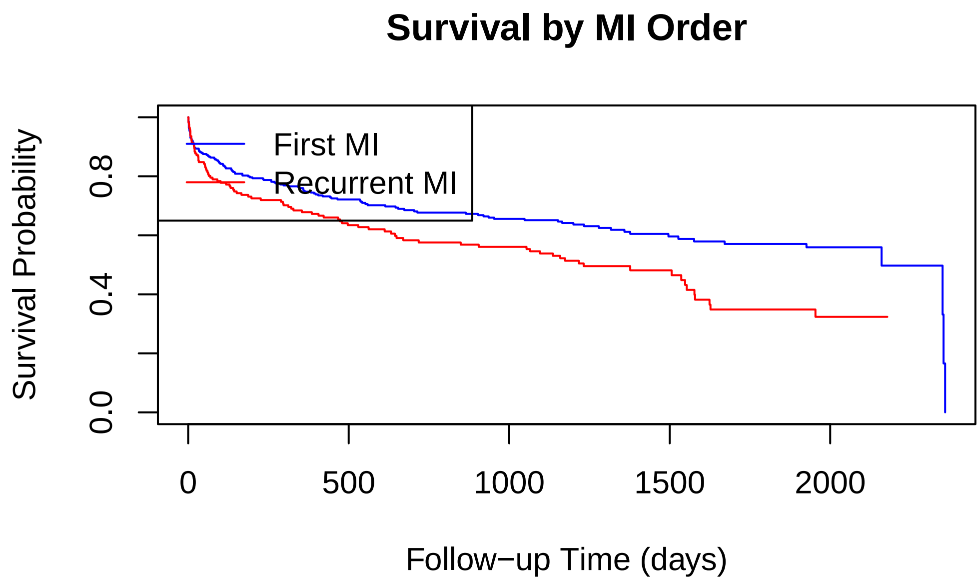
419	223	1	0.732	0.02447	0.6854	0.781
442	215	1	0.728	0.02459	0.6818	0.778
446	210	1	0.725	0.02472	0.6781	0.775
465	202	1	0.721	0.02485	0.6743	0.772
535	186	1	0.718	0.02502	0.6701	0.768
537	185	1	0.714	0.02518	0.6659	0.765
542	184	1	0.710	0.02534	0.6618	0.761
552	181	1	0.706	0.02550	0.6576	0.758
559	179	1	0.702	0.02567	0.6533	0.754
614	170	1	0.698	0.02584	0.6489	0.750
646	168	1	0.694	0.02602	0.6444	0.747
654	167	1	0.689	0.02620	0.6400	0.743
673	164	1	0.685	0.02637	0.6355	0.739
704	162	1	0.681	0.02655	0.6309	0.735
714	161	1	0.677	0.02672	0.6264	0.731
865	159	1	0.673	0.02688	0.6218	0.727
903	158	1	0.668	0.02705	0.6173	0.723
920	157	1	0.664	0.02721	0.6128	0.720
936	156	1	0.660	0.02737	0.6082	0.716
953	155	1	0.655	0.02752	0.6037	0.712
1048	154	1	0.651	0.02767	0.5992	0.708
1152	137	1	0.646	0.02787	0.5941	0.703
1165	132	1	0.642	0.02809	0.5888	0.699
1200	123	1	0.636	0.02834	0.5832	0.694
1233	116	1	0.631	0.02862	0.5772	0.690
1279	105	1	0.625	0.02897	0.5706	0.684
1317	98	1	0.619	0.02937	0.5635	0.679
1359	90	1	0.612	0.02984	0.5559	0.673
1377	87	1	0.605	0.03031	0.5480	0.667
1496	71	1	0.596	0.03106	0.5382	0.660
1527	70	1	0.588	0.03176	0.5285	0.653
1576	69	1	0.579	0.03242	0.5189	0.646
1671	68	1	0.571	0.03304	0.5093	0.639
1926	51	1	0.559	0.03424	0.4961	0.631
2160	9	1	0.497	0.06603	0.3833	0.645
2350	3	1	0.331	0.14230	0.1429	0.769
2353	2	1	0.166	0.13710	0.0328	0.839
2358	1	1	0.000	NaN	NA	NA

miord=recurrent

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
1	171	3	0.982	0.0100	0.963	1.000
2	168	1	0.977	0.0116	0.954	1.000

3	167	2	0.965	0.0141	0.938	0.993
5	165	1	0.959	0.0152	0.930	0.989
6	164	1	0.953	0.0161	0.922	0.985
7	163	3	0.936	0.0188	0.900	0.973
10	160	2	0.924	0.0203	0.885	0.965
11	158	2	0.912	0.0216	0.871	0.956
17	156	1	0.906	0.0223	0.864	0.951
18	155	1	0.901	0.0229	0.857	0.947
19	154	1	0.895	0.0235	0.850	0.942
20	153	2	0.883	0.0246	0.836	0.933
22	151	1	0.877	0.0251	0.829	0.928
26	150	1	0.871	0.0256	0.823	0.923
31	149	1	0.865	0.0261	0.816	0.918
32	148	2	0.854	0.0270	0.802	0.908
33	146	1	0.848	0.0275	0.796	0.904
49	145	1	0.842	0.0279	0.789	0.899
52	144	1	0.836	0.0283	0.783	0.894
53	143	1	0.830	0.0287	0.776	0.889
55	142	1	0.825	0.0291	0.769	0.884
57	141	1	0.819	0.0295	0.763	0.879
60	140	1	0.813	0.0298	0.756	0.873
62	139	1	0.807	0.0302	0.750	0.868
64	138	1	0.801	0.0305	0.744	0.863
69	137	1	0.795	0.0309	0.737	0.858
76	136	1	0.789	0.0312	0.731	0.853
91	135	1	0.784	0.0315	0.724	0.848
101	134	1	0.778	0.0318	0.718	0.843
118	133	1	0.772	0.0321	0.712	0.837
129	132	1	0.766	0.0324	0.705	0.832
132	131	1	0.760	0.0326	0.699	0.827
140	130	1	0.754	0.0329	0.693	0.822
143	129	1	0.749	0.0332	0.686	0.816
151	128	1	0.743	0.0334	0.680	0.811
166	127	1	0.737	0.0337	0.674	0.806
187	126	1	0.731	0.0339	0.667	0.801
197	125	1	0.725	0.0341	0.661	0.795
226	124	1	0.719	0.0344	0.655	0.790
289	123	1	0.713	0.0346	0.649	0.785
295	122	1	0.708	0.0348	0.643	0.779
297	121	1	0.702	0.0350	0.636	0.774
312	120	1	0.696	0.0352	0.630	0.768
321	119	1	0.690	0.0354	0.624	0.763
328	118	1	0.684	0.0355	0.618	0.758

354	117	1	0.678	0.0357	0.612	0.752
385	114	1	0.672	0.0359	0.606	0.747
406	111	1	0.666	0.0361	0.599	0.741
422	109	1	0.660	0.0363	0.593	0.735
467	103	1	0.654	0.0365	0.586	0.729
473	102	1	0.647	0.0367	0.579	0.723
479	100	1	0.641	0.0369	0.573	0.717
497	98	1	0.634	0.0371	0.566	0.711
530	94	1	0.628	0.0373	0.559	0.705
562	87	1	0.620	0.0376	0.551	0.699
612	84	1	0.613	0.0378	0.543	0.692
632	82	1	0.606	0.0381	0.535	0.685
644	81	1	0.598	0.0384	0.527	0.678
649	80	1	0.591	0.0386	0.520	0.671
670	79	1	0.583	0.0388	0.512	0.664
718	78	1	0.576	0.0390	0.504	0.658
849	77	1	0.568	0.0392	0.496	0.651
905	76	1	0.561	0.0394	0.489	0.644
1054	75	1	0.553	0.0396	0.481	0.637
1065	74	1	0.546	0.0398	0.473	0.630
1096	73	1	0.538	0.0399	0.465	0.623
1136	68	1	0.530	0.0401	0.457	0.615
1159	65	1	0.522	0.0403	0.449	0.608
1174	61	1	0.514	0.0406	0.440	0.600
1217	57	1	0.505	0.0408	0.431	0.591
1232	55	1	0.495	0.0411	0.421	0.583
1377	35	1	0.481	0.0423	0.405	0.572
1506	29	1	0.465	0.0440	0.386	0.559
1536	28	1	0.448	0.0454	0.367	0.547
1548	27	1	0.432	0.0467	0.349	0.533
1553	26	1	0.415	0.0477	0.331	0.520
1577	25	1	0.398	0.0486	0.314	0.506
1579	24	1	0.382	0.0494	0.296	0.492
1624	23	1	0.365	0.0499	0.279	0.477
1627	22	1	0.349	0.0503	0.263	0.463
1954	14	1	0.324	0.0525	0.235	0.445



Question 1

Based on the Kaplan-Meier table, what are the estimated survival probabilities at 3 years for patients with a first MI and those with a recurrent MI?

Answer question 1

3 years is equal to 3 * 365 = 1095 days. The estimated survival probabilities at 3 years can be read of the Kaplan-Meier table by finding the closest time point less than or equal to 1095 days for each group. This gives an estimated survival probability of 0.651 for patients with a first MI and 0.546 for patients with a recurrent MI.

Question 2

Based on the Kaplan-Meier curves, do you observe any differences in survival times between patients with a first MI and those with a recurrent MI?

Answer question 2

The Kaplan-Meier curves suggest that patients with a recurrent MI have lower survival probabilities compared to those with a first MI. This indicates a potential difference in survival times between the two groups.

To formally test the difference in survival between the two groups, we can use the logrank test:

```
Call:
survdiff(formula = Surv(lenfol, fstat_numeric) ~ miord, data = whas500)

                  N Observed Expected (O-E)^2/E (O-E)^2/V
miord=first     329      125    146.1      3.04      9.57
miord=recurrent 171       90     68.9      6.43      9.57

 Chisq= 9.6  on 1 degrees of freedom, p= 0.002
```

Question 3

Based on the results of the logrank test, is there a significant difference in overall survival between patients with a first MI and those with a recurrent MI?

Answer question 3

The p-value from the logrank test is a measure of the evidence against the null hypothesis of no difference in survival between the two groups. A small p-value (typically < 0.05) indicates that there is sufficient evidence to reject the null hypothesis and conclude that there is a significant difference in survival between the groups. In this case, the p-value of 0.002 is less than 0.05, suggesting a significant difference in overall survival between patients with a first MI and those with a recurrent MI.

Cox regression

```
Call:
coxph(formula = Surv(lenfol, fstat_numeric) ~ miord, data = whas500)

  n= 500, number of events= 215

                 coef exp(coef) se(coef)     z Pr(>|z|)
miordrecurrent 0.4266    1.5320   0.1391 3.067  0.00216 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

               exp(coef) exp(-coef) lower .95 upper .95
miordrecurrent     1.532     0.6527     1.166     2.012
```

```
Concordance= 0.54  (se = 0.017 )
Likelihood ratio test= 9.14  on 1 df,   p=0.002
Wald test            = 9.41  on 1 df,   p=0.002
Score (logrank) test = 9.55  on 1 df,   p=0.002
```

Question 4

What is the hazard ratio (HR) for patients with a recurrent MI compared to those with a first MI based on the unadjusted Cox regression model?

Answer question 4

The HR for patients with a recurrent MI compared to those with a first MI is 1.53 based on the unadjusted Cox regression model. This means that patients with a recurrent MI have a 53% higher risk of death compared to patients with a first MI.

Question 5

Does the result of the Cox regression model support the findings from the logrank test regarding the association between MI order and overall survival?

Answer question 5

The p-value associated with the miord variable in the Cox regression model is 0.002, which is consistent with the result of the logrank test. Both tests provide evidence of a significant association between MI order and overall survival, indicating that patients with a recurrent MI have a higher risk of death compared to those with a first MI.

```
Call:
coxph(formula = Surv(lenfol, fstat_numeric) ~ miord + age + gender,
    data = whas500)

  n= 500, number of events= 215

                    coef exp(coef)  se(coef)      z Pr(>|z|)
miordrecurrent  0.188124  1.206983  0.139486  1.349    0.177
age             0.066269  1.068514  0.006266 10.576   <2e-16 ***
genderfemale   -0.062241  0.939657  0.140627 -0.443    0.658
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
                exp(coef) exp(-coef) lower .95 upper .95
miordrecurrent     1.2070     0.8285    0.9183     1.586
age                1.0685     0.9359    1.0555     1.082
genderfemale       0.9397     1.0642    0.7133     1.238

Concordance= 0.729  (se = 0.018 )
Likelihood ratio test= 144.2  on 3 df,   p=<2e-16
Wald test            = 119.7  on 3 df,   p=<2e-16
Score (logrank) test = 128.4  on 3 df,   p=<2e-16
```

Question 6

After adjusting for age and gender, what is the hazard ratio (HR) for patients with a recurrent MI compared to those with a first MI? How does this compare to the unadjusted HR? Can you explain the change in the HR after adjusting for these variables?

Answer question 6

After adjusting for gender and age, the HR for patients with a recurrent MI compared to those with a first MI is 1.21, with a corresponding p-value of 0.177. This adjusted HR is lower than the unadjusted HR of 1.53. The change in the HR after adjusting for these variables suggests that gender and age may confound the association between MI order and overall survival. In this case, the mean age of patients with a recurrent MI (73) is higher than that of patients with a first MI (68.2), which could explain the change in the HR after adjusting for age.

Part 2: Unguided exercises

Excercise 1

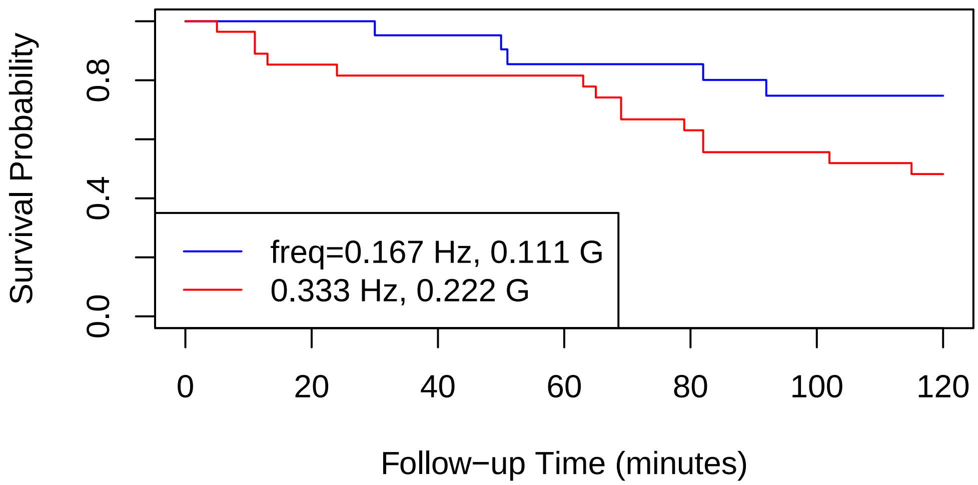
File `Ex9_1.sav` contains data from a small experiment concerning motion sickness at sea (Burns, 1984). Subjects were placed in a cabin subjected to vertical motion for two hours. The outcome variable was the waiting time to emesis (vomiting). Some subjects requested an early stop to the experiment although they had not vomited, yielding censored observations, while others successfully survived two hours. The experiment was carried out with two "treatments": two combinations of movement accelerations and frequency. One combination was used for a group of 21 subjects, the other in a different group of 28 subjects.

```
ex9 <- read_sav("datasets/Ex9_1.sav")
```

```r
# Convert labeled variables to factors
ex9 <- ex9 %>%
  mutate(across(where(is.labelled), as_factor))

# Create a numerical variable for follow-up status (1 = exact, 0 = censored)
ex9$event_numeric <- ifelse(ex9$censor=="exact", 1, 0)
```

(a) Calculate and plot Kaplan-Meier estimates of survival probabilities in the two groups.

```
Call: survfit(formula = Surv(stime, event_numeric) ~ expgroup, data = ex9)

                expgroup=freq=0.167 Hz, 0.111 G
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
   30     21       1    0.952  0.0465        0.866        1.000
   50     20       1    0.905  0.0641        0.788        1.000
   51     18       1    0.854  0.0778        0.715        1.000
   82     16       1    0.801  0.0894        0.644        0.997
   92     15       1    0.748  0.0981        0.578        0.967

                expgroup=freq=0.333 Hz, 0.222 G
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
    5     28       1    0.964  0.0351        0.898        1.000
   11     26       2    0.890  0.0599        0.780        1.000
   13     24       1    0.853  0.0679        0.730        0.997
   24     23       1    0.816  0.0744        0.682        0.976
   63     22       1    0.779  0.0797        0.637        0.952
   65     21       1    0.742  0.0841        0.594        0.926
   69     20       2    0.668  0.0906        0.512        0.871
   79     18       1    0.630  0.0928        0.472        0.841
   82     17       2    0.556  0.0956        0.397        0.779
  102     15       1    0.519  0.0961        0.361        0.746
  115     14       1    0.482  0.0962        0.326        0.713
```



(b) Calculate the 95% CI for the difference between survival probabilities of the two groups after 60 minutes.

According to the previously obtained Kaplan-Meier estimates, the survival probabilities (standard errors) at 60 minutes are 0.854 (0.0778) for the first group and 0.816 (0.0744) for the second group. The estimated difference in survival probabilities is therefore $0.854 - 0.816 = 0.038$. The standard error of the difference is calculated as the square root of the sum of the squared standard errors for each group: $\sqrt{0.0778^2 + 0.0744^2} = 0.108$.

The 95% CI for the difference in survival probabilities after 60 minutes is $0.038 \pm 1.96 \times 0.108 = [-0.174, 0.250]$.

(c) Compare the two survival curves by logrank test.

```
Call:
survdiff(formula = Surv(stime, event_numeric) ~ expgroup, data = ex9)

                                 N Observed Expected (O-E)^2/E (O-E)^2/V
expgroup=freq=0.167 Hz, 0.111 G 21        5     8.86      1.68      3.21
expgroup=freq=0.333 Hz, 0.222 G 28       14    10.14      1.47      3.21

 Chisq= 3.2  on 1 degrees of freedom, p= 0.07
```

The p-value from the logrank test (0.07) is larger than 0.05, suggesting that there is no significant difference in survival probabilities between the two groups.

(d) Use Cox regression to compare the two treatments; compare the result to that of the logrank test; calculate the hazard ratio and its 95% CI.

```
Call:
coxph(formula = Surv(stime, event_numeric) ~ expgroup, data = ex9)

  n= 49, number of events= 19

                                  coef exp(coef) se(coef)     z Pr(>|z|)
expgroupfreq=0.333 Hz, 0.222 G 0.9042    2.4699   0.5214 1.734   0.0829 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                               exp(coef) exp(-coef) lower .95 upper .95
expgroupfreq=0.333 Hz, 0.222 G      2.47     0.4049     0.889     6.862

Concordance= 0.607  (se = 0.054 )
Likelihood ratio test= 3.38  on 1 df,   p=0.07
Wald test            = 3.01  on 1 df,   p=0.08
Score (logrank) test = 3.22  on 1 df,   p=0.07
```

The p-value associated with the `expgroup` variable in the Cox regression model is 0.083, which is consistent with the result of the logrank test. The hazard ratio for the second group compared to the first group is 2.470, with a 95% CI of [0.889, 6.862].

(e) Two different persons undergo the experiment with different treatments. Estimate the probability that the waiting time until emesis under one of the treatments exceeds that under the other treatment.

The probability that the waiting time is shorter under the frequency 0.33 than under the frequency 0.167 is estimated as $2.47/(1 + 2.47) = 0.71$. (calculated using the formula under additional remark 2 in the syllabus).

Exercise 2

Subfertile women with a child wish may receive an in-vitro fertilization (IVF) treatment. In an observational study the waiting time until pregnancy was recorded. The women undergoing the IVF treatment were categorized (prior to the start of the treatment) in four groups, A, B, C and D, with respect to the type of infertility. The waiting times – some of them censored -

of the groups were compared by means of Cox regression with the group variable entered as a categorical variable. The P-value of the Wald test with 3 df was 0.095. The dummy-variables v1, v2 and v3 were defined as follows:

	v1	v2	v3
A	1	0	0
B	0	1	0
C	0	0	1
D	0	0	0

The hazard ratios and P-values for the dummy variables v1, v2 and v3 were reported as 1.20 (P=0.30), 2.80 (P=0.02) and 1.05 (P=0.48).

(a) Which group is estimated to have the longest waiting time for pregnancy?

A HR greater than 1 corresponds to a shorter waiting time. (The event in this exercise is "getting pregnant". A larger hazard means a shorter waiting time). The HR's of A:D, B:D and C:D all exceed 1, and so D has the longest waiting time.

(b) A researcher compared all pairs of groups by separate tests. She used the Bonferroni method to keep the type-I error of the entire procedure below 10%. Which differences were found to be significant?

There are 6 paired comparisons. Thus the Bonferroni corrected P-values are obtained by multiplying the separate P-values by 6. Although not all 6 pairwise P-values are listed, the largest possible HR, 2.8, has a P-value of 0.02, that is 0.12 when corrected. As all the other pairwise p-values are larger, none of the pairwise differences is significant at 5% level.

(c) Which assumptions are needed for the validity of Cox regression in this case?

Non-informative censoring and proportionality of hazard functions

Exercise 3

Relation between survival and a number of variables was studied in 37 patients having a bone marrow transplant. Cox regression analysis using the occurrence of acute graft-versus-host disease (GvHD=1 if present and GvHD=0 if absent), diagnosis, recipient's age and sex, donor's age and sex, whether the donor had been pregnant and the type of the leukemia (CML=1 if chronic myeloid leukemia and CML=0 otherwise) yielded the following model:

Variable	Regression coefficient	Standard error
GvHD(0 = No, 1= Yes)	2.306	0.5898
CML (0 = No, 1 = Yes)	-2.508	0.8095

(a) What is the interpretation of the opposite signs for the regression coefficients?

The opposite signs mean that high values of one variable and low values of the other variable are associated with an increased risk of dying. A positive regression coefficient means that high values of that variable are associated with worse survival, and conversely for a negative coefficient. Thus the model predicts that survival is worse for non-CML patients and those with GvHD.

(b) Calculate the relative risks of dying (hazard ratio) for the following patients relative to non-GvHD non-CML patients (i) with GvHD but not CML, (ii) CML but without GvHD, (iii) CML and GvHD.

We need to calculate the linear predictor for each group of patients. These are as follows:

- non-GvHD & non-CML (reference): 0.000
- GvHD & non-CML: 2.306
- non-GvHD & CML: -2.508
- GvHD & CML: 2.306 + -2.508 = -0.202

The relative risks of dying in the other groups relative to non-GvHD non-CML patients are as follows:

- GvHD & non-CML $\exp(2.306) = 10.03$
- non-GvHD & CML $\exp(-2.508) = 0.08$
- GvHD & CML $\exp(-0.202) = 0.82$