CITY UNIVERSITY OF HONG KONG
STUDENTS' UNION
http://www.cityusu.net/
Name = CHAN King Yeung
SID = 1155119394

STAT 4008

Assignment 2

Question 1

a) $\text{Var}\{\arcsin[\sqrt{\hat{S}(t_0)}]\} = \left[\frac{1}{\sqrt{1-\hat{S}(t_0)}} \cdot \frac{1}{2\sqrt{\hat{S}(t_0)}}\right]^2 \text{Var}[\hat{S}(t_0)]$

$= \frac{\hat{S}(t_0)\,\sigma_S^2(t_0)}{4[1-\hat{S}(t_0)]}$

b) with asymptotical normality, we have $\arcsin[\sqrt{\hat{S}(t_0)}] \sim N\left(\arcsin[\sqrt{S(t_0)}], \frac{S(t_0)\,\sigma_S^2(t_0)}{4[1-S(t_0)]}\right)$ such that

$P\left\{-z_{1-\alpha/2} \le \frac{\arcsin[\sqrt{\hat{S}(t_0)}] - \arcsin[\sqrt{S(t_0)}]}{\frac{1}{2}\sigma_S(t_0)\sqrt{\frac{\hat{S}(t_0)}{1-\hat{S}(t_0)}}} \le z_{1-\alpha/2}\right\} = 1-\alpha$ given that the range of $\arcsin(Y) \in [0, \frac{\pi_0}{2}]$, we have

$\max\left\{0, \arcsin[\sqrt{\hat{S}(t_0)}] - z_{1-\alpha/2}\cdot\frac{1}{2}\sigma_S(t_0)\sqrt{\frac{\hat{S}(t_0)}{1-\hat{S}(t_0)}}\right\} \le \arcsin[\sqrt{S(t_0)}] \le \min\left\{\frac{\pi_0}{2}, \arcsin[\sqrt{\hat{S}(t_0)}] + z_{1-\alpha/2}\cdot\frac{1}{2}\sigma_S(t_0)\sqrt{\frac{\hat{S}(t_0)}{1-\hat{S}(t_0)}}\right\}$

$\Rightarrow \sin^2\left\{\max\left\{0, \arcsin[\sqrt{\hat{S}(t_0)}] - z_{1-\alpha/2}\cdot\frac{1}{2}\sigma_S(t_0)\sqrt{\frac{\hat{S}(t_0)}{1-\hat{S}(t_0)}}\right\}\right\} \le S(t_0) \le \sin^2\left\{\min\left\{\frac{\pi_0}{2}, \arcsin[\sqrt{\hat{S}(t_0)}] + z_{1-\alpha/2}\cdot\frac{1}{2}\sigma_S(t_0)\sqrt{\frac{\hat{S}(t_0)}{1-\hat{S}(t_0)}}\right\}\right\}$

Question 2

a) $L(S) = [S(10-0)-S(10)]^3[S(11-0)-S(11)]^2[S(13-0)-S(13)]^4[S(14-0)-S(14)]^5[S(15-0)-S(15)]^3[S(16-0)-S(16)]^2[S(17-0)-S(17)][S(18-0)-S(18)]$
$\times[S(19-0)-S(19)]^2 S(11)[S(12)]^2 S(13)S(14)[S(16)]^3[S(17)]^2[S(18)]^2[S(19)]^2$

b) please refer to the 5th column

c) please refer to the 6th column

d) $\hat{\mu} = 15.817$

$\text{S.E.}(\hat{\mu}) = \sqrt{\sum\left[S\hat{S}(t_i)\,dt_i\right]^2 \frac{d_i}{n_i(n_i-d_i)}}$

$\approx \sqrt{0.2582}$

≈ 0.5082

t_i	d_i	n_i	λ_i	$\hat{S}_{KM}(t_i)$	$\hat{S}_{NA}(t_i)$	$\int_{t_i}^{t_{max}} \hat{S}_{KM}(t_i)\,dt_i$
0	0	39	0	1	1	15.817
10	3	39	$\frac{1}{13}$	0.9231	0.926	5.817
11	2	36	$\frac{1}{18}$	0.8718	0.8759	4.8939
13	4	31	$\frac{4}{31}$	0.7593	0.7699	3.1503
14	5	26	$\frac{5}{26}$	0.6133	0.6352	2.391
15	3	20	$\frac{3}{20}$	0.5213	0.5467	1.7777
16	2	17	$\frac{2}{17}$	0.46	0.486	1.2564
17	1	12	$\frac{1}{12}$	0.4216	0.4472	0.7964
18	1	9	$\frac{1}{9}$	0.3748	0.4002	0.3748
19	2	4	$\frac{1}{2}$	0.1874	0.2427	0

e) linear method: $0.6133 \pm 1.96\sqrt{0.0067}$

$\approx [0.4529, 0.7737]$

log-log transform: $\left[0.6133^{e^{\frac{1.96\sqrt{0.0177}}{\ln(0.6133)}}}, 0.6133^{e^{\frac{1.96\sqrt{0.0177}}{\ln(0.6133)}}}\right]$

$\approx [0.4345, 0.7507]$

arcsine transform: $\left[\sin^2\left\{\max\left[0, \arcsin(\sqrt{0.6133}) - 1.96\cdot\frac{1}{2}\sqrt{0.0177}\cdot\sqrt{\frac{0.6133}{0.3867}}\right]\right\}, \sin^2\left\{\min\left[\frac{\pi_0}{2}, \arcsin(\sqrt{0.6133}) + 1.96\cdot\frac{1}{2}\sqrt{0.0177}\sqrt{\frac{0.6133}{0.3867}}\right]\right\}\right]$

$\approx [0.4502, 0.7643]$

Name: CHAN King Yeung
SID:1155119394
STAT4008 Assignment 2

Question 3

```
time = c(2, 4, 4, 4, 5, 6, 7, 7, 8)
censor = c(1, 1, 1, 0, 1, 0, 1, 0, 0)
suv = survfit(Surv(time, censor) ~ 1)
summary(suv)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
2	9	1	0.889	0.105	0.706	1.000
4	8	2	0.667	0.157	0.420	1.000
5	5	1	0.533	0.173	0.282	1.000
7	3	1	0.356	0.186	0.128	0.989

```
print(suv, print.rmean = T)
```

The estimated mean is $\hat{\mu} = 5.8667$ and the estimated standard error is $s.e.(\hat{\mu}) = 0.7153$.

Question 4

For the diploid group,

a)

```
timeDT = c(1, 3, 4, 5, 5, 8, 23, 26, 27, 30, 42, 56, 62, 69, 104, 104, 112,
129, 181, 8, 67, 76, 104, 176, 231)
censorDT = c(rep(1, 19), rep(0, 6))
suvDT = survfit(Surv(timeDT, censorDT) ~ 1)
summary(suvDT)
```

time	n.risk	n.event	survival	std.err
1	25	1	0.9600	0.0392
3	24	1	0.9200	0.0543
4	23	1	0.8800	0.0650
5	22	2	0.8000	0.0800
8	20	1	0.7600	0.0854
23	18	1	0.7178	0.0905
26	17	1	0.6756	0.0945
27	16	1	0.6333	0.0976
30	15	1	0.5911	0.0998
42	14	1	0.5489	0.1012
56	13	1	0.5067	0.1019
62	12	1	0.4644	0.1017
69	10	1	0.4180	0.1016
104	8	2	0.3135	0.0995
112	5	1	0.2508	0.0974
129	4	1	0.1881	0.0910
181	2	1	0.0941	0.0806

b)

```
hazardDT = cbind(suvDT$time, suvDT$n.risk, suvDT$n.event, -log(suvDT$surv),
sqrt(cumsum(suvDT$n.event / (suvDT$n.risk * (suvDT$n.risk -
suvDT$n.event)))), suvDT$cumhaz, suvDT$std.chaz)
colnames(hazardDT) = c('time', 'n.risk', 'n.event', 'KM estimate', 's.e. of
KM estimate', 'NA estimate', 's.e. of NA estimate')
hazardDT[hazardDT[, 'n.event'] != 0, ]
```

time	n.risk	n.event	KM estimate	s.e. of KM estimate	NA estimate	s.e. of NA estimate
1	25	1	0.04082199	0.04082483	0.04000000	0.04000000
3	24	1	0.08338161	0.05897678	0.08166667	0.05775908
4	23	1	0.12783337	0.07385489	0.12514493	0.07229433
5	22	2	0.22314355	0.10000000	0.21605402	0.09674038
8	20	1	0.27443685	0.11239030	0.26605402	0.10889767
23	18	1	0.33159526	0.12609343	0.32160957	0.12225024
26	17	1	0.39221988	0.13991434	0.38043310	0.13566624
27	16	1	0.45675840	0.15408663	0.44293310	0.14937061
30	15	1	0.52575127	0.16883304	0.50959977	0.16357269
42	14	1	0.59985925	0.18438845	0.58102834	0.17848827
56	13	1	0.67990195	0.20102079	0.65795142	0.19435849
62	12	1	0.76691333	0.21905505	0.74128475	0.21147025
69	10	1	0.87227385	0.24309715	0.84128475	0.23392236
104	8	2	1.15995592	0.31743171	1.09128475	0.29320585
112	5	1	1.38309947	0.38828197	1.29128475	0.35492206
129	4	1	1.67078154	0.48383492	1.54128475	0.43413093
181	2	1	2.36392872	0.85679416	2.04128475	0.66217042

c)

```
print(suvDT, print.rmean = T)
c(80.2 - 1.96 * 15.6, 80.2 + 1.96 * 15.6)
```

```
      n     events     *rmean *se(rmean)     median    0.95LCL    0.95UCL
   25.0       19.0       80.2       15.6       62.0       27.0      129.0
 * restricted mean with upper limit =  231
```

The estimated mean is $\hat{\mu} = 80.2$ and the 95% CI is $[49.624, 110.776]$.

d)

```
medianDT = cbind(summary(suvDT)$time, (summary(suvDT)$surv - 0.5) /
summary(suvDT)$std.err)
colnames(medianDT) = c('time', 'linear')
medianDT
```

time	linear
1	11.73713835
3	7.74070270
4	5.84684582
5	3.75000000
8	3.04390389
23	2.40619654
26	1.85733939
27	1.36628540
30	0.91294536
42	0.48304992
56	0.06545539
62	-0.34947847
69	-0.80697058
104	-1.87409231
112	-2.55901762
129	-3.42712049
181	-5.03775741

The estimated median is 62 and the 95% CI is $[26, 112)$.

For aneuploid group,

a)

```
timeAT = c(1, 3, 3, 4, 10, 13, 13, 16, 16, 24, 26, 27, 28, 30, 30, 32, 41,
51, 65, 67, 70, 72, 73, 77, 91, 93, 96, 100, 104, 157, 167, 61, 74, 79, 80,
81, 87, 87, 88, 89, 97, 101, 104, 108, 109, 120, 131, 150, 231, 240, 400)
censorAT = c(rep(1, 31), rep(0, 20))
suvAT = survfit(Surv(timeAT, censorAT) ~ 1)
summary(suvAT)
```

time	n.risk	n.event	survival	std.err
1	51	1	0.980	0.0194
3	50	2	0.941	0.0329
4	48	1	0.922	0.0376
10	47	1	0.902	0.0416
13	46	2	0.863	0.0482
16	44	2	0.824	0.0534
24	42	1	0.804	0.0556
26	41	1	0.784	0.0576
27	40	1	0.765	0.0594
28	39	1	0.745	0.0610
30	38	2	0.706	0.0638
32	36	1	0.686	0.0650
41	35	1	0.667	0.0660
51	34	1	0.647	0.0669
65	32	1	0.627	0.0678
67	31	1	0.607	0.0686
70	30	1	0.586	0.0692
72	29	1	0.566	0.0697
73	28	1	0.546	0.0701
77	26	1	0.525	0.0705
91	18	1	0.496	0.0723
93	17	1	0.467	0.0737
96	16	1	0.437	0.0747
100	14	1	0.406	0.0756
104	12	1	0.372	0.0765
157	5	1	0.298	0.0905
167	4	1	0.223	0.0936

b)

```
hazardAT = cbind(suvAT$time, suvAT$n.risk, suvAT$n.event, -log(suvAT$surv),
sqrt(cumsum(suvAT$n.event / (suvAT$n.risk * (suvAT$n.risk -
suvAT$n.event)))), suvAT$cumhaz, suvAT$std.chaz)
colnames(hazardAT) = c('time', 'n.risk', 'n.event', 'KM estimate', 's.e. of
KM estimate', 'NA estimate', 's.e. of NA estimate')
hazardAT[hazardAT[, 'n.event'] != 0, ]
```

time	n.risk	n.event	KM estimate	s.e. of KM estimate	NA estimate	s.e. of NA estimate
1	51	1	0.01980263	0.01980295	0.01960784	0.01960784
3	50	2	0.06062462	0.03500700	0.05960784	0.03441609
4	48	1	0.08167803	0.04085037	0.08044118	0.04023053
10	47	1	0.10318424	0.04616587	0.10171777	0.04551032
13	46	2	0.14763600	0.05585185	0.14519603	0.05492147
16	44	2	0.19415601	0.06482037	0.19065058	0.06363510
24	42	1	0.21825357	0.06915490	0.21446010	0.06794350
26	41	1	0.24294618	0.07343131	0.23885035	0.07218867
27	40	1	0.26826399	0.07767356	0.26385035	0.07639505
28	39	1	0.29423947	0.08190205	0.28949137	0.08058329
30	38	2	0.34830669	0.09038769	0.34212295	0.08876208
32	36	1	0.37647757	0.09467621	0.36990073	0.09300706
41	35	1	0.40546511	0.09901475	0.39847216	0.09729665
51	34	1	0.43531807	0.10341754	0.42788392	0.10164493
65	32	1	0.46706677	0.10818157	0.45913392	0.10634027
67	31	1	0.49985659	0.11304212	0.49139199	0.11112532
70	30	1	0.53375814	0.11801672	0.52472532	0.11601701
72	29	1	0.56884946	0.12312381	0.55920808	0.12103309
73	28	1	0.60521711	0.12838311	0.59492236	0.12619238
77	26	1	0.64443782	0.13424115	0.63338390	0.13192349
91	18	1	0.70159624	0.14590634	0.68893946	0.14314408
93	17	1	0.76222086	0.15800358	0.74776299	0.15475928
96	16	1	0.82675938	0.17068039	0.81026299	0.16690322
100	14	1	0.90086735	0.18608144	0.88169156	0.18154538
104	12	1	0.98787873	0.20543140	0.96502489	0.19975778
157	5	1	1.21102228	0.30364792	1.16502489	0.28267149
167	4	1	1.49870435	0.41896944	1.41502489	0.37736345

c)

```
print(suvAT, print.rmean = T)
c(144.2 - 1.96 * 27.3, 144.2 + 1.96 * 27.3)
```

n	events	*rmean	*se(rmean)	median	0.95LCL	0.95UCL
51.0	31.0	144.2	27.3	91.0	67.0	NA

The estimated mean is $\hat{\mu} = 144.2$ and the 95% CI is $[90.692, 197.708]$.

d)

```
medianAT = cbind(summary(suvAT)$time, (summary(suvAT)$surv - 0.5) /
summary(suvAT)$std.err)
colnames(medianAT) = c('time', 'linear')
medianAT
```

time	linear
1	24.74378710
3	13.39017830
4	11.19810727
10	9.65328284
13	7.52803203
16	6.06070482
24	5.46669517
26	4.93658651
27	4.45652072
28	4.01635069
30	3.22684056
32	2.86691410
41	2.52487623
51	2.19762268
65	1.87042991
67	1.55479726
70	1.24842840
72	0.94931367
73	0.65565541
77	0.35415436
91	-0.05815271
93	-0.45261730
96	-0.83752733
100	-1.24068626
104	-1.66851728
157	-2.23435516
167	-2.95473355

The estimated median is 91 and the 95% CI is $[65, 157)$.

Question 5

a)

```
data = read.csv('ass2q5.csv')
data$Censor = ifelse(data$Censor == 0, 1, 0)
suvQ5 = survfit(Surv(data$Time, data$Censor) ~ 1)
summary(suvQ5)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0.0	149	1	0.9933	0.00669	0.98026	1.0000
0.3	148	1	0.9866	0.00943	0.96827	1.0000
1.1	147	2	0.9732	0.01324	0.94754	0.9995
1.5	145	1	0.9664	0.01475	0.93796	0.9958
1.8	144	1	0.9597	0.01611	0.92868	0.9918
1.9	143	1	0.9530	0.01733	0.91964	0.9876
2.5	140	1	0.9462	0.01850	0.91064	0.9832
3.0	139	1	0.9394	0.01958	0.90181	0.9786
3.5	138	1	0.9326	0.02059	0.89311	0.9738
3.6	137	2	0.9190	0.02242	0.87607	0.9640
4.5	133	1	0.9121	0.02330	0.86754	0.9589
4.6	132	1	0.9052	0.02412	0.85910	0.9537
4.7	131	1	0.8983	0.02491	0.85074	0.9484
5.0	129	1	0.8913	0.02567	0.84237	0.9431
5.5	128	1	0.8843	0.02640	0.83408	0.9376
5.8	125	1	0.8773	0.02712	0.82568	0.9320
5.9	124	1	0.8702	0.02781	0.81735	0.9264
6.7	122	1	0.8630	0.02848	0.80900	0.9207
6.8	120	2	0.8487	0.02976	0.79229	0.9091
7.2	117	3	0.8269	0.03154	0.76734	0.8911
7.3	113	2	0.8123	0.03264	0.75075	0.8788
8.1	109	1	0.8048	0.03318	0.74235	0.8725
8.5	108	1	0.7974	0.03370	0.73398	0.8662
8.6	107	1	0.7899	0.03420	0.72565	0.8599
8.9	106	1	0.7825	0.03468	0.71736	0.8535
9.0	105	1	0.7750	0.03514	0.70911	0.8470
9.6	104	1	0.7676	0.03558	0.70089	0.8406
10.2	101	5	0.7296	0.03766	0.65936	0.8072
10.3	96	3	0.7068	0.03871	0.63481	0.7869
10.5	93	2	0.6916	0.03935	0.61859	0.7731
10.6	91	2	0.6764	0.03992	0.60247	0.7593
10.8	88	1	0.6687	0.04020	0.59435	0.7523
10.9	86	1	0.6609	0.04048	0.58614	0.7452
11.0	84	1	0.6530	0.04075	0.57785	0.7380
11.1	82	3	0.6291	0.04153	0.55278	0.7160
11.2	79	1	0.6212	0.04176	0.54449	0.7087
11.3	78	1	0.6132	0.04198	0.53621	0.7013
11.4	77	2	0.5973	0.04237	0.51975	0.6864
11.5	75	1	0.5893	0.04255	0.51156	0.6789
11.6	74	2	0.5734	0.04286	0.49524	0.6639
11.7	72	1	0.5654	0.04300	0.48713	0.6563
11.9	71	1	0.5575	0.04313	0.47903	0.6487
12.0	70	1	0.5495	0.04324	0.47096	0.6411
12.1	69	1	0.5415	0.04334	0.46292	0.6335
12.3	67	2	0.5254	0.04353	0.44662	0.6180
12.4	65	13	0.4203	0.04350	0.34313	0.5148
12.5	52	2	0.4041	0.04330	0.32758	0.4986
12.6	49	1	0.3959	0.04320	0.31966	0.4903
12.9	48	1	0.3876	0.04308	0.31177	0.4820
13.0	47	1	0.3794	0.04294	0.30391	0.4736
13.4	46	2	0.3629	0.04263	0.28826	0.4568
13.5	44	1	0.3546	0.04245	0.28048	0.4484
13.6	43	1	0.3464	0.04226	0.27273	0.4400
13.7	41	1	0.3380	0.04206	0.26479	0.4313
13.8	39	1	0.3293	0.04187	0.25665	0.4225
13.9	38	1	0.3206	0.04165	0.24855	0.4136
14.0	37	3	0.2946	0.04089	0.22446	0.3867
14.1	34	4	0.2600	0.03958	0.19289	0.3504
14.2	30	2	0.2426	0.03879	0.17736	0.3319
14.3	28	4	0.2080	0.03692	0.14685	0.2945
14.4	23	2	0.1899	0.03586	0.13115	0.2749
14.5	21	4	0.1537	0.03328	0.10057	0.2350
14.6	17	1	0.1447	0.03252	0.09312	0.2248
14.8	16	2	0.1266	0.03087	0.07849	0.2042
15.0	14	1	0.1175	0.02996	0.07133	0.1937
15.4	12	4	0.0784	0.02559	0.04132	0.1486
15.5	8	1	0.0686	0.02419	0.03434	0.1369
15.6	7	1	0.0588	0.02263	0.02763	0.1250
15.7	6	1	0.0490	0.02087	0.02124	0.1129
15.8	5	2	0.0294	0.01649	0.00978	0.0883
16.3	3	1	0.0196	0.01360	0.00503	0.0763
16.5	2	1	0.0098	0.00971	0.00141	0.0683
16.9	1	1	0.0000	NaN	NA	NA

b)

```
suv5 = summary(suvQ5)
suv5 = data.frame(suv5$time, suv5$n.risk, suv5$n.event, suv5$surv,
suv5$std.err)
colnames(suv5) = c('time', 'n.risk', 'n.event', 'survival', 'std.err')
suv5[suv5$time == 5,]
```

	time	n.risk	n.event	survival	std.err
5	5	129	1	0.8912915	0.02566962

c)

```
print(suvQ5, print.rmean = T)
medianQ5 = cbind(summary(suvQ5)$time, (summary(suvQ5)$surv - 0.5) /
summary(suvQ5)$std.err)
colnames(medianQ5) = c('time', 'linear')
medianQ5
```

n	events	*rmean	*se(rmean)	median	0.95LCL	0.95UCL
149.000	124.000	11.289	0.329	12.400	11.600	12.500

* restricted mean with upper limit = 16.9

time	linear
0.0	73.7478928
0.3	51.6128066
1.1	35.7328992
1.5	31.6163471
1.8	28.5456582
1.9	26.1338693
2.5	24.1207020
3.0	22.4432431
3.5	21.0141151
3.6	18.6847353
4.5	17.6889957
4.6	16.7964696
4.7	15.9891387
5.0	15.2433733
5.5	14.5597587
5.8	13.9121156
5.9	13.3130237
6.7	12.7483165
6.8	11.7141013
7.2	10.3642134
7.3	9.5679361
8.1	9.1874691
8.5	8.8247262
8.6	8.4779890
8.9	8.1457571
9.0	7.8267131
9.6	7.5196943
10.2	6.0956651
10.3	5.3406184
10.5	4.8686743
10.6	4.4176827
10.8	4.1957981
10.9	3.9749424
11.0	3.7550051
11.1	3.1894253
11.2	2.9015453
11.3	2.6068375
11.4	2.2959041
11.5	2.0991990
11.6	1.7122101
11.7	1.5215181
11.9	1.3324376
12.0	1.1447856
12.1	0.9583860
12.3	0.5828885
12.4	-1.8323183
12.5	-2.2139865
12.6	-2.4103238
12.9	-2.6084981
13.0	-2.8087873
13.4	-3.2162345
13.5	-3.4240448
13.6	-3.6349186
13.7	-3.8525831
13.8	-4.0775424
13.9	-4.3065973
14.0	-5.8226542
14.1	-6.0643815
14.2	-6.6342394
14.3	-7.9096455
14.4	-8.6487822
14.5	-10.4063834
14.6	-10.9258372
14.8	-12.8961497
15.0	-12.7653587
15.4	-16.4769144
15.5	-17.8327844
15.6	-19.4945156
15.7	-21.6073472
15.8	-28.5351164
16.3	-35.3338247
16.5	-50.5090758
16.9	NaN

The estimated median is 12.4 and the 95% CI is $[11.6, 12.5)$.

Question 6

```
timeS = c(1, 6, 5, 3, 7, 2, 3, 4, 5)
timeE = c(1, 6, 5, 3, 9, 2, 4, 4, 5)
type = c(1, 0, 2, 1, 3, 2, 3, 1, 0)
suvQ6 = survfit(Surv(timeS, timeE, type, type = 'interval') ~ 1)
summary(suvQ6)
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
1.0	9.0	2.399363	0.733	0.1178	0.477	1.00
1.5	6.6	0.000653	0.733	0.1178	0.477	1.00
3.0	6.6	1.199995	0.600	0.1031	0.342	1.00
4.0	5.4	2.399989	0.333	0.0538	0.129	0.86
8.0	1.0	1.000000	0.000	0.0000	NA	NA

Question 7

```
entry = c(58, 58, 59, 60, 60, 61, 61, 62 ,62, 62, 63, 63, 64, 66, 66, 67, 67,
67, 68, 69, 69, 69, 70, 70, 70, 71, 72, 72, 73, 73)
exit = c(60, 63, 69, 62, 65, 72, 69, 73, 66, 65, 68, 74, 71, 68, 69, 70, 77,
69, 72, 79, 72, 70, 76, 71, 78, 79, 76, 73, 80, 74)
cens = c(1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0,
1, 0, 0, 1, 1, 0, 1)
data = data.frame(entry, exit, cens)

suvQ7 = survfit(Surv(exit, cens) ~ 1, data[exit >= 60, ])
ti = suvQ7$time[suvQ7$n.event != 0]
ni = rep(0, length(ti))
for(i in 1:length(ti))
  ni[i] = sum(entry <= ti[i] & ti[i]<= exit)
di = suvQ7$n.event[suvQ7$n.event != 0]
surv_fun_b = cumprod(1 - di / ni)

suvQ7 = survfit(Surv(exit, cens) ~ 1, data[exit >= 65, ])
ti_c = suvQ7$time[suvQ7$n.event != 0]
ni_c = rep(0, length(ti_c))
for(i in 1:length(ti_c))
  ni_c[i] = sum(entry <= ti_c[i] & ti_c[i]<= exit)
di_c = suvQ7$n.event[suvQ7$n.event != 0]
surv_fun_c = cumprod(1 - di_c / ni_c)

Q7 = data.frame(ti, ni, round(surv_fun_b, 4), round(c(rep(1, length(ti) -
length(ti_c)), surv_fun_c), 4))
colnames(Q7) = c('Time', '  Y', 'S(t|60)', 'S(t|65)')
Q7
```

Time	Y	S(t|60)	S(t|65)
60	5	0.8000	1.0000
62	9	0.7111	1.0000
63	10	0.6400	1.0000
65	10	0.5120	0.8000
66	10	0.4608	0.7200
68	13	0.3899	0.6092
69	14	0.3342	0.5222
70	13	0.2828	0.4419
71	12	0.2357	0.3682
72	12	0.1964	0.3068
73	11	0.1785	0.2790
74	9	0.1587	0.2480
76	7	0.1360	0.2125
77	5	0.1088	0.1700

a) Please refer to the 2nd column.

b) Please refer to the 3rd column.

c) Please refer to the 4th column.

Question 8

```
ti = c(2, 4, 7, 14, 20, 18, 8, 13, 17, 26, 20, 15, 23, 5, 16, 15, 11, 9, 33, 4,
8, 35, 10, 36, 25)
xi = c(30, 27, 25, 19, 18, 17, 16, 16, 15, 15, 15, 13, 13, 12, 11, 9, 8, 8, 8,
7, 6, 6, 6, 4, 4)
tau = 42
ri = tau - xi
suvQ8 = survfit(Surv(ri) ~ 1)
temp1 = outer(ti, suvQ8$time, '<=')
temp2 = outer(ri, suvQ8$time, '>=')
temp3 = temp1 & temp2
ni = apply(temp3, 2, sum)
surv = cumprod(1 - suvQ8$n.event / ni)

di = rep(0, length(ri))
temp4 = outer(ri,ri,'==')
di[1] = sum(temp4[, 1])
for(i in 2:length(ri)) {
  if(all(temp4[, i] == temp4[, i - 1]))
    di[i] = 0
  else
    di[i] = sum(temp4[, i])
}

ni = rep(0, length(ri))
for(i in 1:length(ri))
  ni[i] = sum(ri[i] <= ri & ri[i] >= ti)
ni[which(di == 0)] = 0

surv_fun = rep(0, length(ri))
surv_fun[which(di != 0)] = round(surv, 4)

surv_fun[which(di == 0)] = ''
ni[which(di == 0)] = ''
di = ifelse(di == 0, '', di)
Q8 = data.frame(ti, xi, ri, di, ni, surv_fun)
colnames(Q8) = c('  T_i', '  X_i', '  R_i', '  d_i', '  n_i', 'Pr(X < x_i|X <=
42)')
Q8
```

T_i	X_i	R_i	d_i	n_i	Pr(X < x_i|X <= 42)
2	30	12	1	10	0.9
4	27	15	1	13	0.8308
7	25	17	1	14	0.7714
14	19	23	1	17	0.7261
20	18	24	1	16	0.6807
18	17	25	1	16	0.6381
8	16	26	2	16	0.5584
13	16	26			
17	15	27	3	14	0.4387
26	15	27			
20	15	27			
15	13	29	2	11	0.3589
23	13	29			
5	12	30	1	9	0.3191
16	11	31	1	8	0.2792
15	9	33	1	8	0.2443
11	8	34	3	7	0.1396
9	8	34			
33	8	34			
4	7	35	1	5	0.1117
8	6	36	3	5	0.0447
35	6	36			
10	6	36			
36	4	38	2	2	0
25	4	38			