Introduction to TTE modeling: Workbook 2

Kaplan-Meier estimates and plots

2023-07-11

Contents

Preliminaries for R examples

```
library(tidyverse)
library(stringr)
library(survival)
library(survminer)
library(texreg)
library(mgcv)
library(flexsurv)
library(muhaz)
library(Hmisc)

theme_set(theme_bw())
```

Estimating the survival function

We'll use the `survfit` function to estimate the Survival function.

Using our example data from the last class, let's estimate the overall S(t), i.e., not stratifying by any covariates.

```
load('../data/aedat.RDS')

aedat <-
  aedat %>%
  mutate(AETOXGR = factor(aedat$AETOXGR, 0:3, labels=c("None","Mild","Moderate","Severe")),
         ae_any = AETOXGR != 'None') %>%
  group_by(USUBJID) %>%
  # End of study for patients without a severe event
  mutate(TTE_SEVERE = case_when(
    STUDYID=="PROTA" ~ 2,
    STUDYID=="PROTB" ~ 6
  ),
  # Time of severe event for those that had one
  TTE_SEVERE = ifelse(AETOXGR=="Severe", TTE, TTE_SEVERE)
  )
```

```r
# Both for EDA and for model-checking, it's generally helpful to have quartiles of exposure:
dat_use <-
  aedat %>% arrange(USUBJID, TTE_SEVERE) %>% slice(1) %>%
  group_by(PBO) %>%
  mutate(Quartile = ifelse(PBO == "PBO", "PBO",
                           paste0("Q", ntile(CAVGSS, n = 4)))) %>%
  ungroup() %>%
  mutate(rowid = 1:n())
```

The key elements:

- `survfit`: the function used to obtain the K-M (or Flemming-Harrington) estimate of the survival function
  - stype=1 (the default) gives the K-M estimate
  - stype=2 gives the Fleming-Harrington estimate
- `Surv` function for defining the outcome
  - For right-censored data, the first argument is the observed event time variable, and the second argument is the even indicator (1 or TRUE for an event; 0 or FALSE for censoring)

Obtaining the estimate and looking at some output:

```r
km_est = survfit(Surv(TTE,ae_any)~1, data = dat_use)

print(km_est)
```

```
. Call: survfit(formula = Surv(TTE, ae_any) ~ 1, data = dat_use)
.
.        n events median 0.95LCL 0.95UCL
. [1,] 180    132   0.32   0.139   0.649
```

```r
summary(km_est)
```

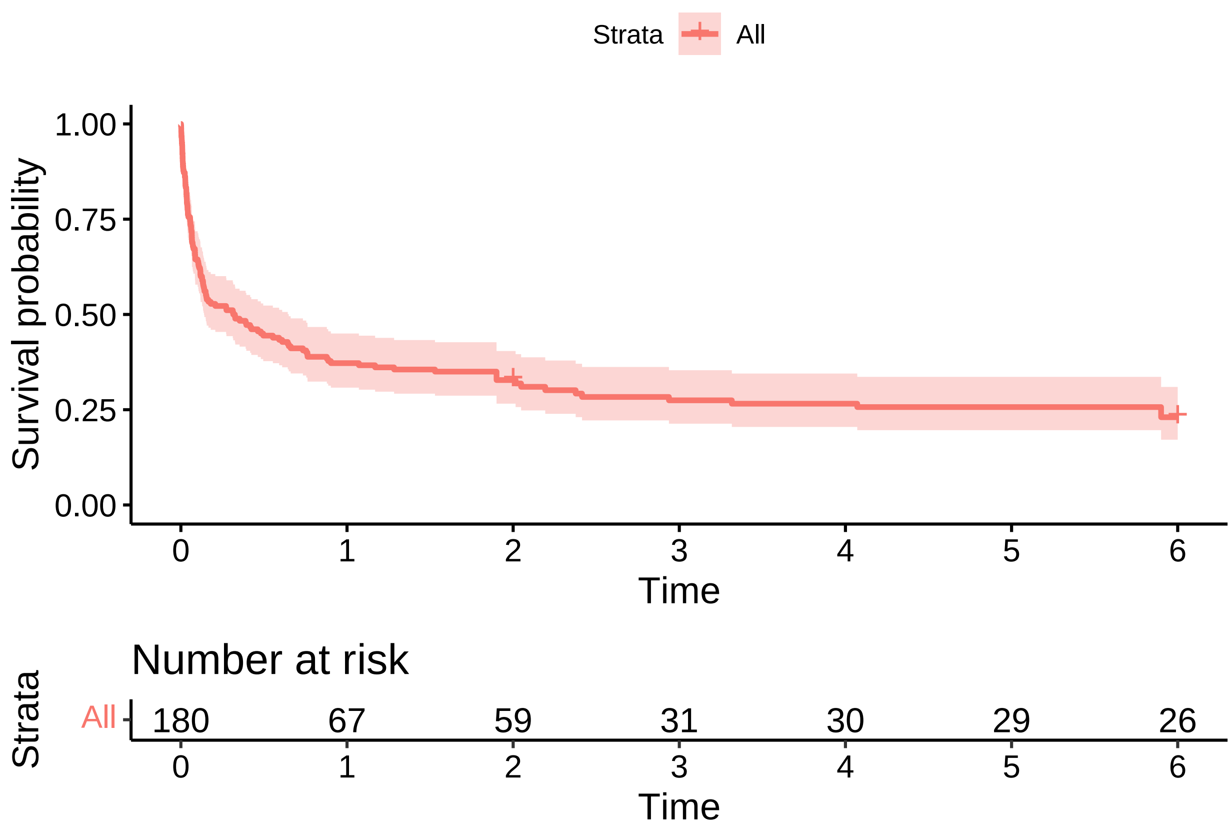
```
. Call: survfit(formula = Surv(TTE, ae_any) ~ 1, data = dat_use)
.
.      time n.risk n.event survival std.err lower 95% CI upper 95% CI
.  0.000805    180       1    0.994 0.00554        0.984        1.000
.  0.001935    179       1    0.989 0.00781        0.974        1.000
.  0.002245    178       1    0.983 0.00954        0.965        1.000
.  0.002302    177       1    0.978 0.01099        0.956        1.000
.  0.003077    176       1    0.972 0.01225        0.949        0.997
.  0.003199    175       1    0.967 0.01338        0.941        0.993
.  0.004871    174       1    0.961 0.01441        0.933        0.990
.  0.005289    173       1    0.956 0.01536        0.926        0.986
.  0.006043    172       1    0.950 0.01624        0.919        0.982
.  0.006828    171       1    0.944 0.01707        0.912        0.979
.  0.008177    170       1    0.939 0.01785        0.905        0.975
.  0.008267    169       1    0.933 0.01859        0.898        0.970
.  0.008707    168       1    0.928 0.01929        0.891        0.966
.  0.008826    167       1    0.922 0.01996        0.884        0.962
.  0.010071    166       1    0.917 0.02060        0.877        0.958
.  0.010159    165       1    0.911 0.02121        0.870        0.954
.  0.010784    164       1    0.906 0.02180        0.864        0.949
.  0.010916    163       1    0.900 0.02236        0.857        0.945
.  0.011968    162       1    0.894 0.02290        0.851        0.940
.  0.012724    161       1    0.889 0.02342        0.844        0.936
.  0.013258    160       1    0.883 0.02393        0.838        0.931
```

.	0.014633	159	1	0.878	0.02441	0.831	0.927
.	0.016949	158	1	0.872	0.02488	0.825	0.922
.	0.023699	157	1	0.867	0.02534	0.818	0.918
.	0.023706	156	1	0.861	0.02578	0.812	0.913
.	0.026148	155	1	0.856	0.02620	0.806	0.908
.	0.026399	154	1	0.850	0.02661	0.799	0.904
.	0.026941	153	1	0.844	0.02701	0.793	0.899
.	0.027425	152	1	0.839	0.02740	0.787	0.894
.	0.028619	151	1	0.833	0.02778	0.781	0.890
.	0.032233	150	1	0.828	0.02814	0.774	0.885
.	0.032706	149	1	0.822	0.02850	0.768	0.880
.	0.032818	148	1	0.817	0.02884	0.762	0.875
.	0.033797	147	1	0.811	0.02917	0.756	0.870
.	0.034673	146	1	0.806	0.02950	0.750	0.865
.	0.035457	145	1	0.800	0.02981	0.744	0.861
.	0.035762	144	1	0.794	0.03012	0.738	0.856
.	0.036722	143	1	0.789	0.03042	0.731	0.851
.	0.039233	142	1	0.783	0.03071	0.725	0.846
.	0.039720	141	1	0.778	0.03099	0.719	0.841
.	0.040568	140	1	0.772	0.03126	0.713	0.836
.	0.041805	139	1	0.767	0.03153	0.707	0.831
.	0.042022	138	1	0.761	0.03178	0.701	0.826
.	0.045127	137	1	0.756	0.03203	0.695	0.821
.	0.055222	136	1	0.750	0.03227	0.689	0.816
.	0.055578	135	1	0.744	0.03251	0.683	0.811
.	0.057453	134	1	0.739	0.03274	0.677	0.806
.	0.057879	133	1	0.733	0.03296	0.671	0.801
.	0.061409	132	1	0.728	0.03318	0.666	0.796
.	0.061810	131	1	0.722	0.03338	0.660	0.791
.	0.063174	130	1	0.717	0.03359	0.654	0.786
.	0.065723	129	1	0.711	0.03378	0.648	0.781
.	0.065771	128	1	0.706	0.03397	0.642	0.775
.	0.066141	127	1	0.700	0.03416	0.636	0.770
.	0.066437	126	1	0.694	0.03433	0.630	0.765
.	0.066679	125	1	0.689	0.03451	0.624	0.760
.	0.070688	124	1	0.683	0.03467	0.619	0.755
.	0.072054	123	1	0.678	0.03483	0.613	0.750
.	0.075189	122	1	0.672	0.03499	0.607	0.744
.	0.084249	121	1	0.667	0.03514	0.601	0.739
.	0.084579	120	1	0.661	0.03528	0.595	0.734
.	0.084672	119	1	0.656	0.03542	0.590	0.729
.	0.086239	118	1	0.650	0.03555	0.584	0.724
.	0.086305	117	1	0.644	0.03568	0.578	0.718
.	0.101183	116	1	0.639	0.03580	0.572	0.713
.	0.104850	115	1	0.633	0.03592	0.567	0.708
.	0.106140	114	1	0.628	0.03603	0.561	0.703
.	0.109378	113	1	0.622	0.03614	0.555	0.697
.	0.115346	112	1	0.617	0.03624	0.550	0.692
.	0.117250	111	1	0.611	0.03634	0.544	0.687
.	0.117528	110	1	0.606	0.03643	0.538	0.681
.	0.118619	109	1	0.600	0.03651	0.533	0.676
.	0.127254	108	1	0.594	0.03660	0.527	0.671
.	0.127997	107	1	0.589	0.03667	0.521	0.665
.	0.133053	106	1	0.583	0.03675	0.516	0.660

```
. 0.133380   105       1    0.578 0.03681        0.510         0.655
. 0.136157   104       1    0.572 0.03688        0.504         0.649
. 0.139387   103       1    0.567 0.03694        0.499         0.644
. 0.141296   102       1    0.561 0.03699        0.493         0.638
. 0.149190   101       1    0.556 0.03704        0.488         0.633
. 0.150323   100       1    0.550 0.03708        0.482         0.628
. 0.153820    99       1    0.544 0.03712        0.476         0.622
. 0.155550    98       1    0.539 0.03715        0.471         0.617
. 0.165857    97       1    0.533 0.03718        0.465         0.611
. 0.179546    96       1    0.528 0.03721        0.460         0.606
. 0.207173    95       1    0.522 0.03723        0.454         0.601
. 0.272220    94       1    0.517 0.03725        0.449         0.595
. 0.273933    93       1    0.511 0.03726        0.443         0.590
. 0.312540    92       1    0.506 0.03727        0.438         0.584
. 0.314848    91       1    0.500 0.03727        0.432         0.579
. 0.325156    90       1    0.494 0.03727        0.427         0.573
. 0.327124    89       1    0.489 0.03726        0.421         0.568
. 0.353456    88       1    0.483 0.03725        0.416         0.562
. 0.389702    87       1    0.478 0.03723        0.410         0.557
. 0.393778    86       1    0.472 0.03721        0.405         0.551
. 0.416938    85       1    0.467 0.03718        0.399         0.546
. 0.423652    84       1    0.461 0.03715        0.394         0.540
. 0.462944    83       1    0.456 0.03712        0.388         0.534
. 0.481351    82       1    0.450 0.03708        0.383         0.529
. 0.495438    81       1    0.444 0.03704        0.377         0.523
. 0.553784    80       1    0.439 0.03699        0.372         0.518
. 0.591294    79       1    0.433 0.03694        0.367         0.512
. 0.609558    78       1    0.428 0.03688        0.361         0.507
. 0.643238    77       1    0.422 0.03681        0.356         0.501
. 0.649290    76       1    0.417 0.03675        0.351         0.495
. 0.660878    75       1    0.411 0.03667        0.345         0.490
. 0.734580    74       1    0.406 0.03660        0.340         0.484
. 0.754551    73       1    0.400 0.03651        0.334         0.478
. 0.761845    72       1    0.394 0.03643        0.329         0.473
. 0.762268    71       1    0.389 0.03634        0.324         0.467
. 0.878613    70       1    0.383 0.03624        0.318         0.461
. 0.886580    69       1    0.378 0.03614        0.313         0.456
. 0.902642    68       1    0.372 0.03603        0.308         0.450
. 1.071367    67       1    0.367 0.03592        0.303         0.444
. 1.169048    66       1    0.361 0.03580        0.297         0.439
. 1.283269    65       1    0.356 0.03568        0.292         0.433
. 1.529470    64       1    0.350 0.03555        0.287         0.427
. 1.900000    63       4    0.328 0.03499        0.266         0.404
. 2.014711    37       1    0.319 0.03515        0.257         0.396
. 2.047774    36       1    0.310 0.03527        0.248         0.387
. 2.192981    35       1    0.301 0.03536        0.239         0.379
. 2.376302    34       1    0.292 0.03541        0.231         0.371
. 2.414482    33       1    0.283 0.03543        0.222         0.362
. 2.937613    32       1    0.275 0.03541        0.213         0.354
. 3.316241    31       1    0.266 0.03536        0.205         0.345
. 4.071150    30       1    0.257 0.03527        0.196         0.336
. 5.900000    29       3    0.230 0.03480        0.171         0.310
```

Plotting the estimate using `ggsurvplot`:

```
ggsurvplot(km_est, risk.table = TRUE)
```



Often it's helpful to add the number of subjects at risk to the bottom of the plot (a risk table).

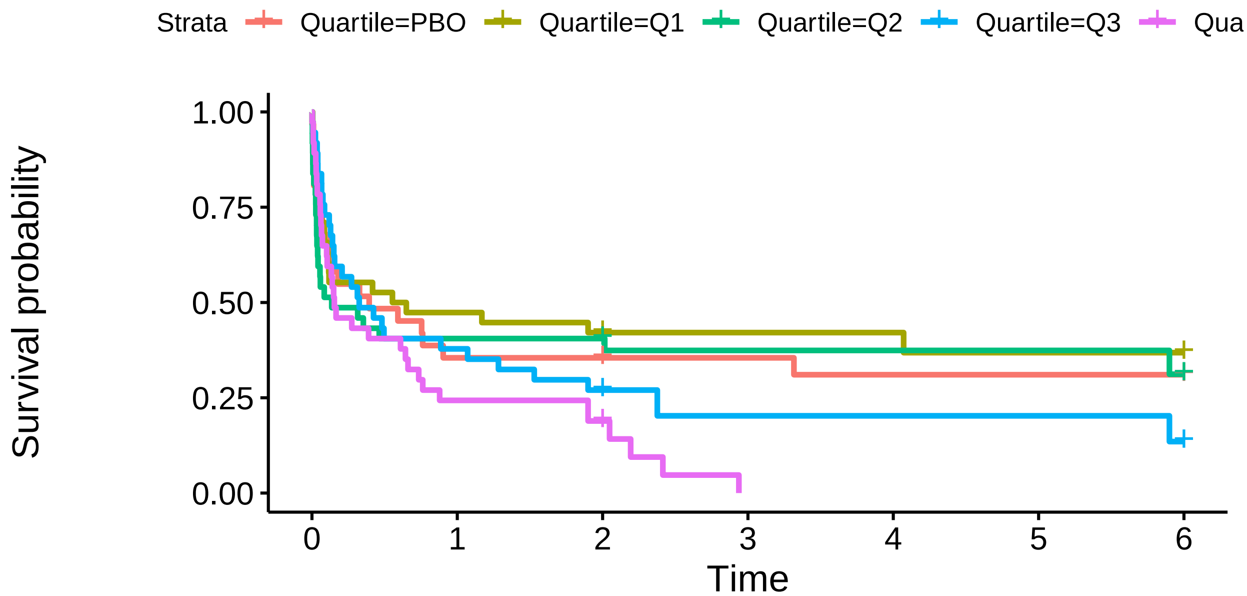
Exercise:

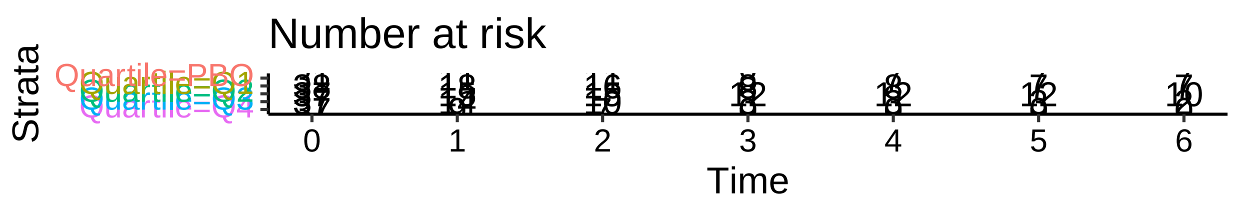
1. Obtain the Kaplan-Meier estimates for time to any AE by exposure quartile. Hint: the right hand side of the formula will be "~ Quartile"

```
km_est_css = survfit(Surv(TTE,ae_any)~Quartile, data = dat_use)
```

2. Plot the estimates using ggsurvplot. Do you see any evidence of an exposure-response relationship?

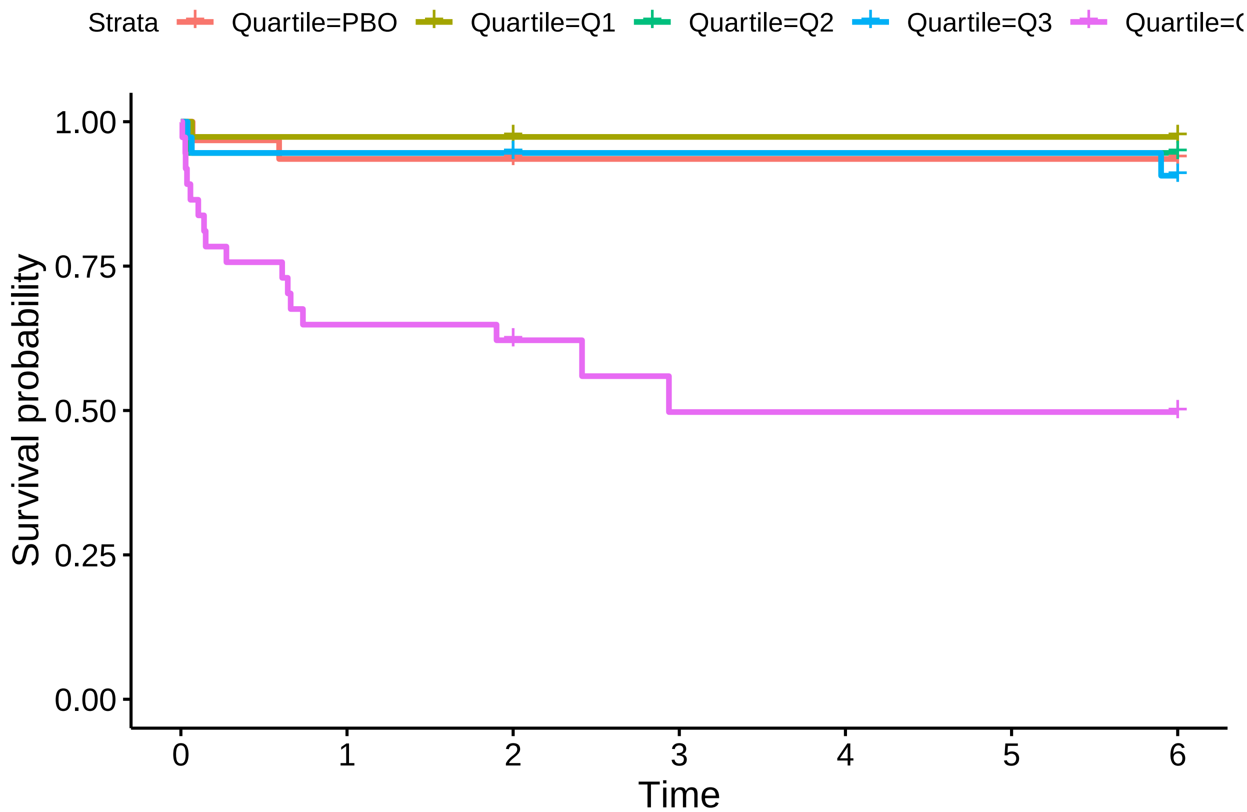
```
ggsurvplot(km_est_css, risk.table = TRUE)
```





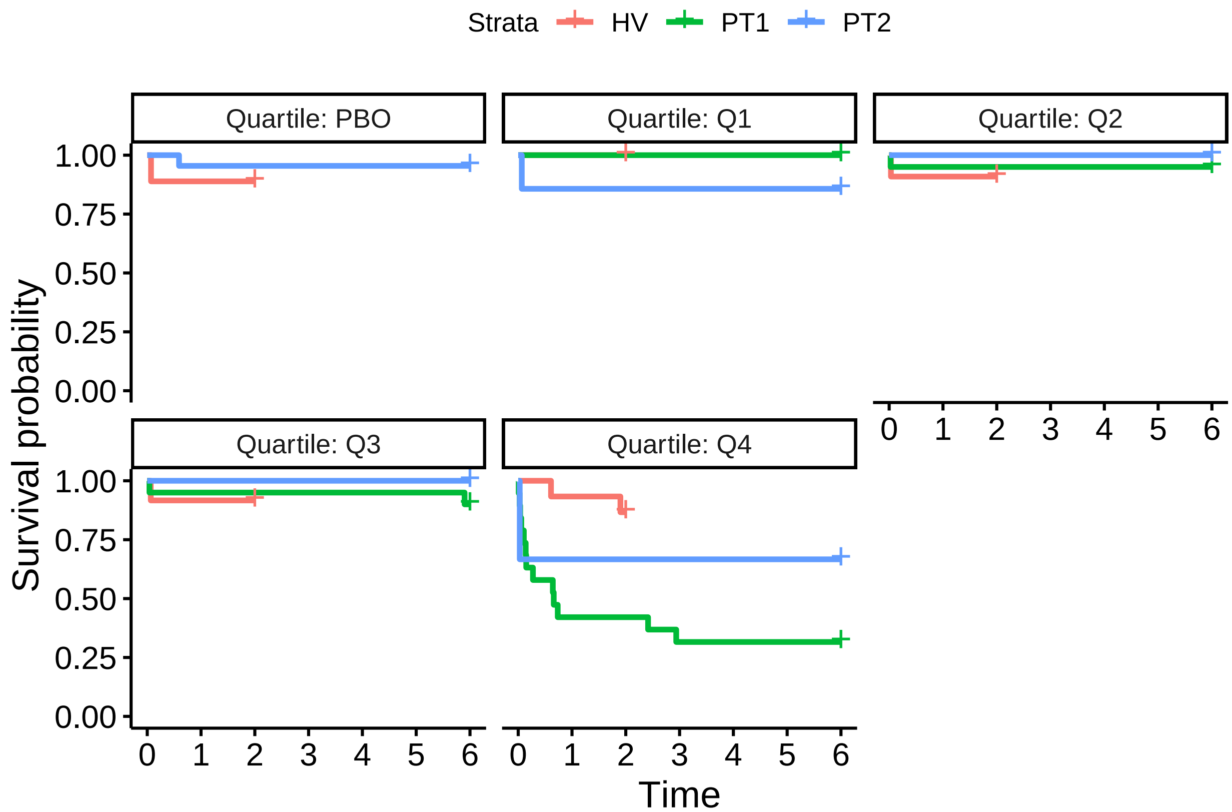
3. Repeat steps 1 and 2 for time to a severe AE (event time variable is TTE_SEVERE, and the event indicator variable is AE01). Do you see any evidence of an exposure-response relationship?

```
km_est_severe_css = survfit(Surv(TTE_SEVERE,AE01)~Quartile, data = dat_use)
ggsurvplot(km_est_severe_css)
```



3. Extra: Plot the K-M estimates by patient type, faceted by exposure quartile. Hint: Fit model `Surv(TTE,AY_any)~PTTYPE` and use the `facet.by` argument to `ggsurvplot`

```
fit <- survfit(Surv(TTE_SEVERE,AE01)~PTTYPE, data=dat_use)
ggsurvplot(fit, facet.by = 'Quartile', data=dat_use)
```



Strata
HV
PT1
PT2
Quartile: PBO
Quartile: Q1
Quartile: Q2
Quartile: Q3
Quartile: Q4
Survival probability
1.00
0.75
0.50
0.25
0.00
0 1 2 3 4 5 6
Time