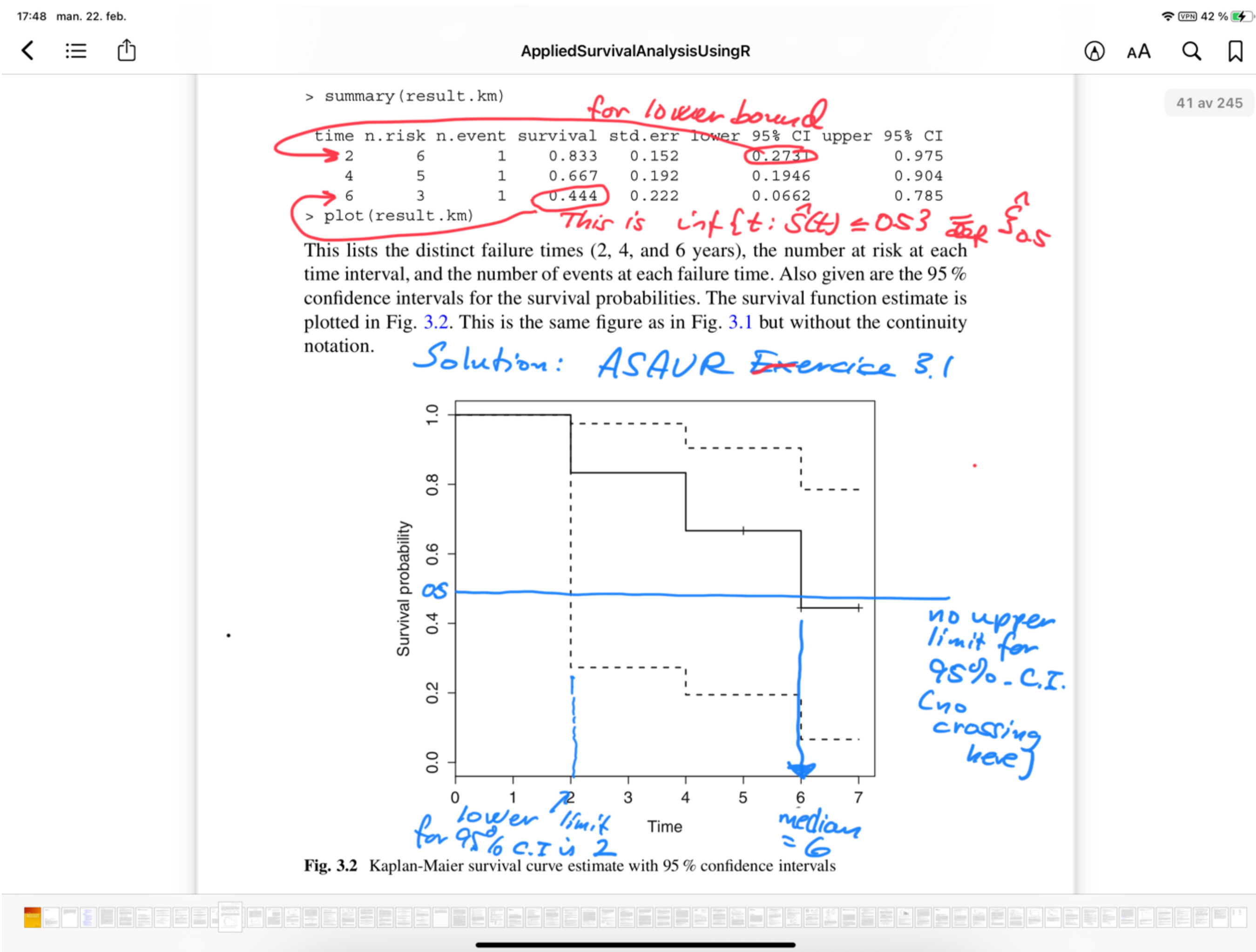
# Solution ASAUR Exercise 3.1

See Solution of Exercise 3.2 below for more information.



```
> summary(result.km)
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
    2      6       1    0.833   0.152       0.2731        0.975
    4      5       1    0.667   0.192       0.1946        0.904
    6      3       1    0.444   0.222       0.0662        0.785
> plot(result.km)
```

This lists the distinct failure times (2, 4, and 6 years), the number at risk at each time interval, and the number of events at each failure time. Also given are the 95 % confidence intervals for the survival probabilities. The survival function estimate is plotted in Fig. 3.2. This is the same figure as in Fig. 3.1 but without the continuity notation.



Fig. 3.2 Kaplan-Maier survival curve estimate with 95 % confidence intervals

# Solution ASAUR Exercise 3.2

Let `result.km` be as derived from the R-code on p. 30 of ASAUR.

```
> summary(result.km)
Call: survfit(formula = Surv(timeMonths, delta) ~ 1, conf.type = "log-log")

  time n.risk n.event survival std.err lower 95% CI upper 95% CI
 0.926     48       1    0.979  0.0206        0.861        0.997
 1.851     47       3    0.917  0.0399        0.793        0.968
 2.083     44       1    0.896  0.0441        0.768        0.955
 2.545     43       1    0.875  0.0477        0.743        0.942
 2.777     42       1    0.854  0.0509        0.718        0.928
```

```
  3.008     41       1   0.833  0.0538         0.694        0.913
  3.702     40       2   0.792  0.0586         0.647        0.882
  3.934     38       2   0.750  0.0625         0.602        0.850
  4.397     36       1   0.729  0.0641         0.580        0.833
  4.860     35       1   0.708  0.0656         0.558        0.816
  5.554     34       2   0.667  0.0680         0.515        0.781
  5.785     32       1   0.646  0.0690         0.494        0.763
  6.479     31       2   0.604  0.0706         0.452        0.726
  6.942     29       1   0.583  0.0712         0.432        0.708
  8.562     28       2   0.542  0.0719         0.392        0.670
  9.719     26       1   0.521  0.0721         0.372        0.650
  9.950     25       1   0.500  0.0722         0.353        0.631
 10.645     23       1   0.478  0.0722         0.332        0.610
 12.264     19       1   0.453  0.0727         0.308        0.587
 13.653     16       1   0.425  0.0735         0.280        0.562
 13.884     14       1   0.394  0.0742         0.251        0.535
 14.810     13       1   0.364  0.0744         0.223        0.507
 15.273     12       1   0.334  0.0742         0.196        0.478
 17.587     11       1   0.303  0.0734         0.170        0.449
 18.050     10       1   0.273  0.0720         0.145        0.418
> quantile(result.km,probs=c(.25,.5,.75))
$quantile
       25        50        75
 4.165289 10.297521        NA

$lower
       25        50        75
 2.545455  5.785124 14.809917

$upper
       25        50        75
 6.479339 15.272727        NA
```

Thus there are two R-commands used here,

```
summary(result.km)
quantile(result.km,probs=c(.25,.5,.75))
```

Now recall the definition

$$\hat{\xi}_p = \inf\{t : \hat{S}(t) \leq 1 - p\}$$

Quantiles can then either be found by the above `quantile` command, or by looking at the above `summary` output. For example, for estimation of the lower quartile, we should then look for the smallest $t$ for which $\hat{S}(t) \leq 0.75$. Graphically, we then go horizontally at height 0.5 until we reach the curve - usually a vertical line, which then goes down to the desired value. Otherwise, we could use the summary output from R. Generally, we should then go down to the first time the survival probability goes below 0.75. In the present case, we obtain in fact the *exact* value of 0.75 for $t = 3.934$. Then since $\hat{S}(t)$ is a step function, the estimate 0.75 is valid for all $t$ until $t = 4.397$. It turns out that R then chooses the average of these two numers, namely 4.165 in this case, see the figure below.

(This situation also applies to the estimation of the median in the present case where R estimates the median to be the average $(9.950 + 10.645)/2 = 10.298$).

The figure below shows the graphical solution.

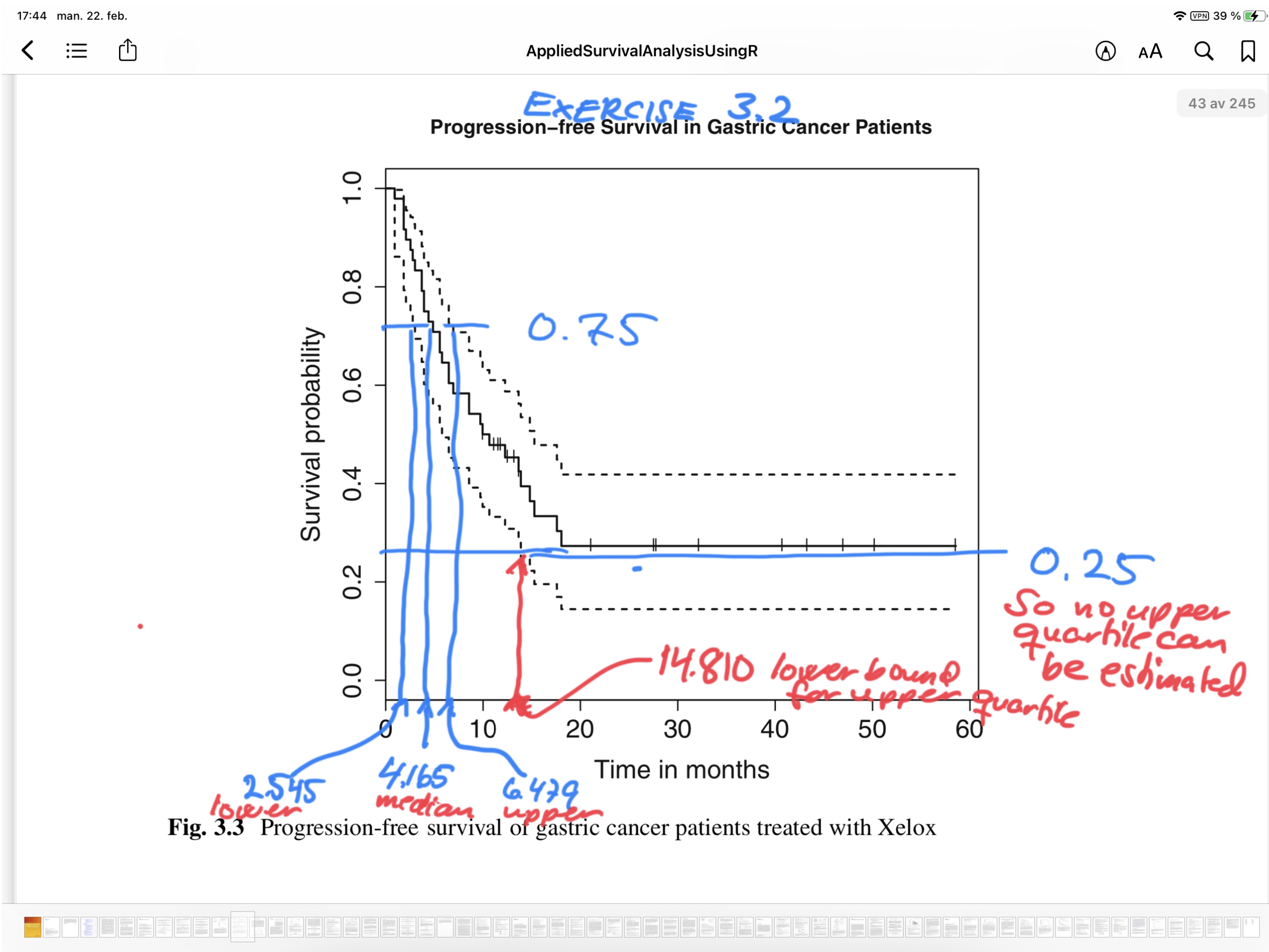



**Fig. 3.3** Progression-free survival of gastric cancer patients treated with Xelox