The predictive individual effect distribution for survival data: Appendix

Beat Neuenschwander, Satrajit Roychoudhury

2021-12-14

Contents

The appendix provides some methodological details and explains the code that has been used for the three examples.

1 The Lehmann-Doksum model

To derive the predictive distribution of the individual effect, we use a model introduced by Lehmann (Lehmann 1974) and investigated by Doksum (Doksum 1974). The Lehmann-Doksum model relates the outcomes X and Y as follows:

$$\text{Lehmann-Doksum model:}$$
$$X + \Delta(X) \text{ has the same distribution as } Y \tag{1}$$

Because the shift function is not assumed constant, this looks like an assumption-free construction. Unfortunately, however, $\Delta(X)$ is not unique. A simple example is $X \sim N(0,1)$ and $Y \sim N(\delta,1)$: besides the natural $\Delta_1(X) = \delta$, $\Delta_2(X) = -2X + \delta$ also satisfies the above requirement.

Doksum defines the *shift function* as the minimum horizontal distance between F and G (see Figure 1) and gives the solution for $\Delta(x)$:
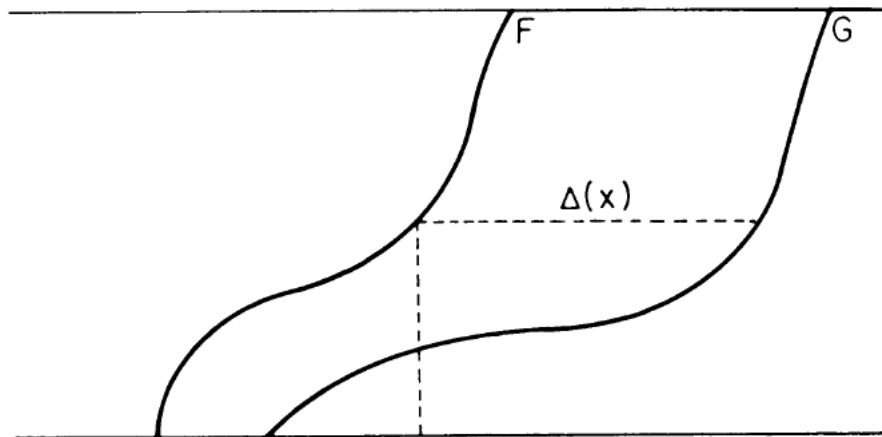
$$\Delta(x) = G^{-1}\{F(x)\} - x \tag{2}$$

Figure 1: The Doksum distance (Doksum 1974)

2 Implementation

2.1 Main function: pie_surv

The main function is **pie_surv**. For help, see the file **pie_surv.r**.

The data set passed to **pie_surv** must contain the following variables:

- time: observed survival or censoring times
- event: event indicator, 1 for survival and 0 for censoring times
- trt: treatment indicator, 1 for treatment and 0 for control

2.2 Required libraries

In addition to the main library **clinpredict**, the following libraries are needed: **survival, ggplot2, tidyverse, mvtnorm, survminer**

```r
# required libraries
library(clinpredict)


library(ggplot2)
library(survival)
library(tidyverse)
library(mvtnorm)
library(survminer)

options(digits=2)

# Simulation sample size for example
# 1000 sample is too small: default is 10000
n.pred = 2000
```

3   Example 1: CM141 trial

3.1   Analysis

```r
# Example data file
data("CM141.r")
ds = CM141

ds = data(CM141)
ds = CM141


C.label = "standard of care"
T.label = "nivolumab"

yC.cond = 1:15
gain.cutoffs = -1:12

gain.label = "survival gain in months (nivolumab-standard)"

out1 = pie_surv(
  data = ds,
  n.pred = n.pred,
  yC.cond = yC.cond,
  pie.gain.cutoffs = gain.cutoffs,
  time.unit = "months",
  gain.label = gain.label,
  C.label = C.label,
  T.label = T.label
)
```

3.2 Outputs

3.2.1 Predictive individual effect distribution

```
# output object has the following elements...
names(out1)
```

```
.. [1] "rank.pres"   "int"        "coxPH"        "logrank"
.. [5] "pwexp"       "sim"        "marg"         "pie"
.. [9] "pie.gain"    "pie.cond"   "tables_knitr" "figures"
```

```
# 1) main results
out1$pie
```

```
..      mean  sd  2.5% 50% 97.5%   <0  >=0
.. [1,]    4 5.3 -0.56 2.5    19 0.17 0.83
```

```
out1$pie.gain
```

```
..   <=-1   <=0    >0    >1    >2    >3    >4    >5    >6    >7    >8    >9
.. 0.008 0.168 0.832 0.619 0.538 0.466 0.391 0.315 0.257 0.205 0.167 0.138
..   >10   >11   >12
.. 0.112 0.085 0.066
```

```
# conditional pie (for fixed value control times)
out1$pie.cond
```

```
..                       mean   sd  2.5%   50% 97.5%     <0  >=0
.. standard of care=1  -0.0914 0.33 -0.62 -0.13  0.62 0.6425 0.36
.. standard of care=2  -0.0043 0.29 -0.49 -0.03  0.69 0.5500 0.45
.. standard of care=3   0.2352 0.73 -0.90  0.19  1.88 0.3995 0.60
.. standard of care=4   1.0046 1.03 -0.89  0.95  3.22 0.1635 0.84
.. standard of care=5   2.8785 1.32  0.57  2.77  5.84 0.0050 0.99
.. standard of care=6   4.7524 1.79  1.59  4.61  8.73 0.0000 1.00
.. standard of care=7   5.5472 2.01  2.12  5.35 10.02 0.0000 1.00
.. standard of care=8   6.0476 2.22  2.27  5.89 10.98 0.0000 1.00
.. standard of care=9   6.5480 2.48  2.32  6.34 12.12 0.0005 1.00
.. standard of care=10  7.0485 2.79  2.36  6.76 13.28 0.0010 1.00
.. standard of care=11  7.5489 3.13  2.33  7.23 14.44 0.0010 1.00
.. standard of care=12  8.0493 3.49  2.31  7.68 15.63 0.0010 1.00
.. standard of care=13  8.5498 3.87  2.19  8.15 16.97 0.0015 1.00
.. standard of care=14  9.0502 4.26  2.04  8.60 18.52 0.0030 1.00
.. standard of care=15  9.5506 4.65  1.88  9.03 19.72 0.0055 0.99
```

3.2.2 Additional output

```
# interval boundaries
out1$int
```

```
.. $intC
.. 10% 50% 80%
.. 1.1 3.9 6.2
..
.. $intT
.. 20% 40%
.. 1.6 2.4
```

```r
# Cox and log-rank results
out1$coxPH
```

```
.. Call:
.. coxph(formula = Surv(time, event) ~ trt, data = data)
..
..      coef exp(coef) se(coef)  z      p
.. trt1 -0.4      0.7      0.1 -3 0.007
..
.. Likelihood ratio test=7  on 1 df, p=0.008
.. n= 361, number of events= 218
```

```r
out1$logrank
```

```
.. [1] 0.0034
```

```r
# summaries  of marginal predictive distributions
out1$marg
```

```
..                  mean   sd 2.5% 50% 97.5%
.. standard of care  7.8  8.4 0.38 4.9    31
.. nivolumab        11.9 12.5 0.37 7.8    47
```

```r
# results from pwexp analysis for control and treatment
# out1$pwexp

# simulated predictive values
head(out1$sim)
```

```
.. [1] "simulated values not saved: set *save.sim=TRUE*"
```

3.2.3 Tables and figures

```r
# knitr tables: out1$tables_knitr
names(out1$tables_knitr)
```

```
.. [1] "marg"     "pie"     "pie.cond" "pie.gain"
```

```r
out1$tables_knitr$marg
```

Table 1: : predictive standard of care and nivolumab survival times

	mean	sd	2.5%	50%	97.5%
standard of care	7.8	8.4	0.38	4.9	31
nivolumab	11.8	12.5	0.37	7.8	47

```
out1$tables_knitr$pie
```

Table 2: : predictive individual effect (T-C)

mean	sd	2.5%	50%	97.5%	<0	>=0
4	5.3	-0.56	2.5	19	0.17	0.83

```
out1$tables_knitr$pie.gain
```

Table 3: : cumulative predictive individual effect distribution

	x
<=-1	0.01
<=0	0.17
>0	0.83
>1	0.62
>2	0.54
>3	0.47
>4	0.39
>5	0.32
>6	0.26
>7	0.21
>8	0.17
>9	0.14
>10	0.11
>11	0.08
>12	0.07

```
# ggplot figures: out1$figures
names(out1$figures)
```

```
.. [1] "km"       "km.pwexp" "pie.gain" "pie.cond"
```

```
fig1.km = out1$figures$km
fig1.km.pwexp = out1$figures$km.pwexp
fig1.gain = out1$figures$pie.gain
fig1.cond = out1$figures$pie.cond

pdf("appfig1km.pdf",onefile=FALSE)
fig1.km
dev.off()
```

```
.. pdf
..    2
```

```
pdf("appfig1kmpwexp.pdf",onefile=FALSE)
fig1.km.pwexp
dev.off()
```
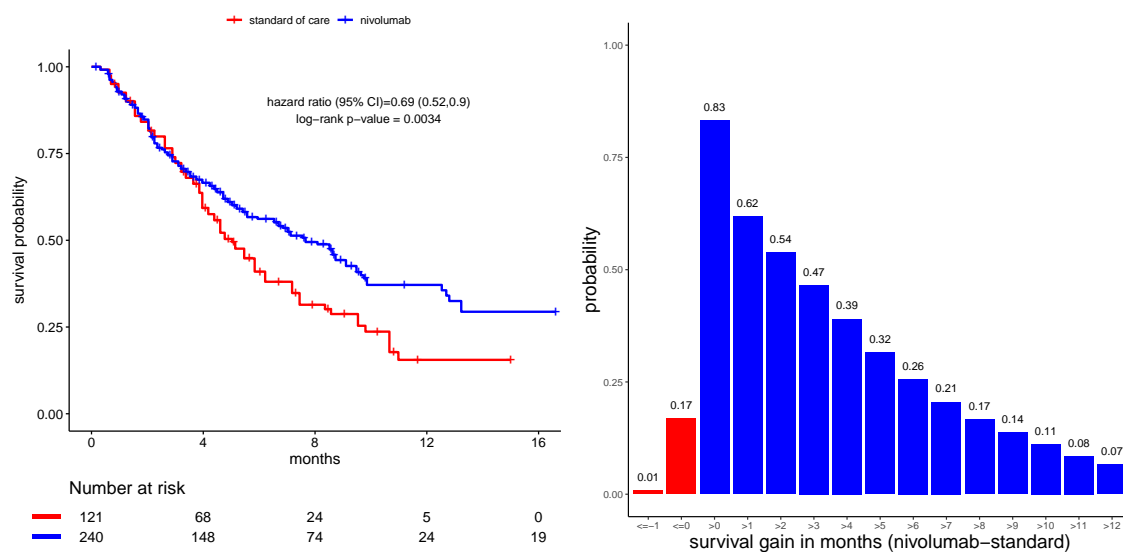
```
.. pdf
..    2
```

```
pdf("appfig1gain.pdf",onefile=FALSE)
fig1.gain
dev.off()
```

```
.. pdf
..    2
```

```
pdf("appfig1cond.pdf",onefile=FALSE)
fig1.cond
dev.off()
```
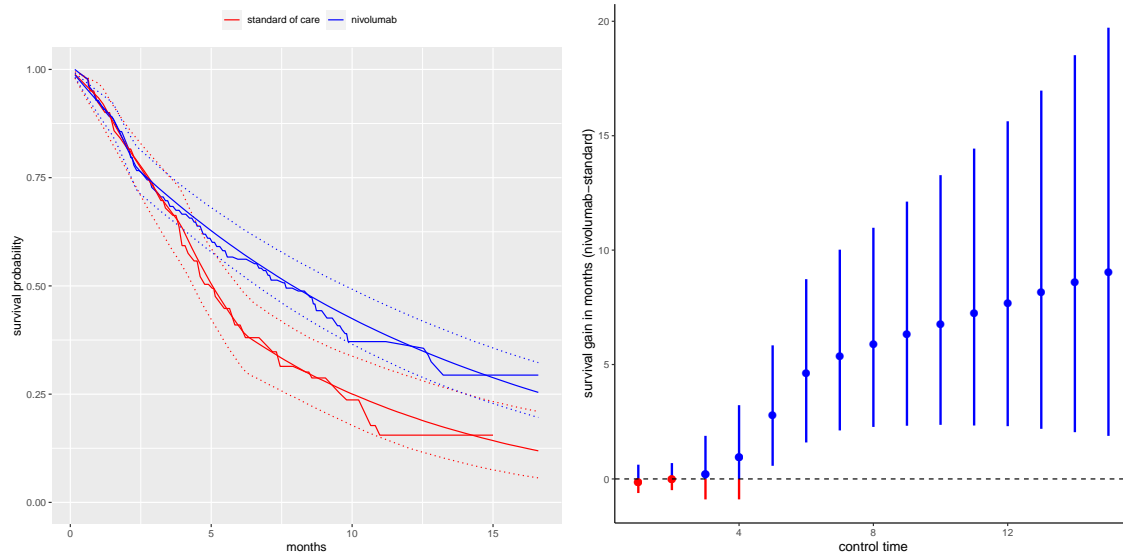
```
.. pdf
..    2
```

Figure 2: Kaplan-Meier curves, piecewise exponential distributions, survival gain, and conditional survival gain for CM141 trial

4 Example 2: CM057 trial with differential treatment effect for biomarker subgroup

The following analysis shows the predictive individual effect distribution for biomarker-negative and biomarker-positive patients in the CM057 trial (Borghaei et al. 2015).

```r
data(CM057)
ds = CM057

# biomarker: group variable

# biomarker-negative
ds0 = ds[ds$group == 0, ]
# biomarker-positive
ds1 = ds[ds$group == 1, ]

C.label = "docetaxel"
T.label = "nivolumab"

gain.cutoffs = seq(-3, 24, by = 3)
gain.label = "survival gain in months (nivolumab-docetaxel)"

out2.0 = pie_surv(
  ds0,
  n.pred = n.pred,
  C.label = C.label,
  T.label = T.label,
  gain.label = gain.label,
  pie.gain.cutoffs = gain.cutoffs
)

out2.1 = pie_surv(
  ds1,
```

```
  n.pred = n.pred,
  C.label = C.label,
  T.label = T.label,
  gain.label = gain.label,
  pie.gain.cutoffs = gain.cutoffs
)

labels = c("biomarker-negative","biomarker-positive")

out.pie = rbind(out2.0$pie,out2.1$pie)
rownames(out.pie) = labels

out.piegain = rbind(out2.0$pie.gain,out2.1$pie.gain)
rownames(out.piegain) = labels

cap = "Summary of predictive individual effect distribution"
knitr::kable(out.pie,cap=cap)
```

Table 4: Summary of predictive individual effect distribution

	mean	sd	2.5%	50%	97.5%	<0	>=0
biomarker-negative	-0.56	4.2	-10.14	-0.48	5.9	0.62	0.38
biomarker-positive	10.60	9.8	-0.86	11.10	30.3	0.08	0.92

```
cap = "Probabilities of survival gains for nivolumab vs. docetaxel"
knitr::kable(rbind(out.piegain),cap=cap)
```

Table 5: Probabilities of survival gains for nivolumab vs. docetaxel

	<=-3	<=0	>0	>3	>6	>9	>12	>15	>18	>21	>24
biomarker-negative	0.10	0.62	0.38	0.11	0.02	0.0	0.00	0.00	0.00	0.00	0.00
biomarker-positive	0.01	0.08	0.92	0.76	0.69	0.6	0.43	0.25	0.15	0.09	0.06

```
fig20.km = out2.0$figures$km
fig21.km = out2.1$figures$km

fig20.gain = out2.0$figures$pie.gain
fig21.gain = out2.1$figures$pie.gain


dev.off()


.. null device
..           1

pdf("appfig20km.pdf",onefile=FALSE)
fig20.km
dev.off()
```

```
.. null device
..            1
```

```
pdf("appfig21km.pdf",onefile=FALSE)
fig21.km
dev.off()
```

```
.. null device
..            1
```
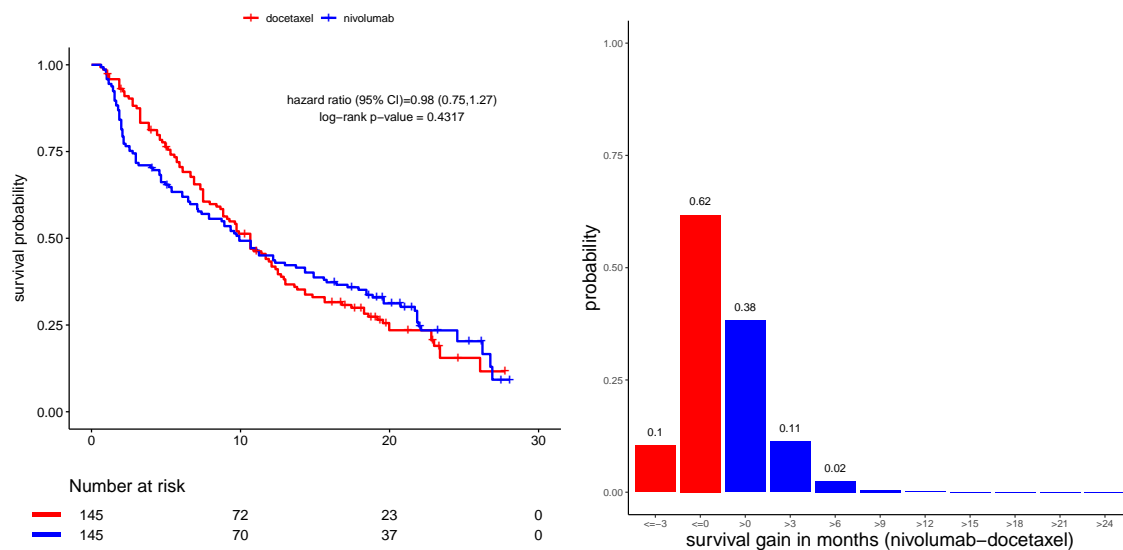
```
pdf("appfig20gain.pdf",onefile=FALSE)
fig20.gain
dev.off()
```
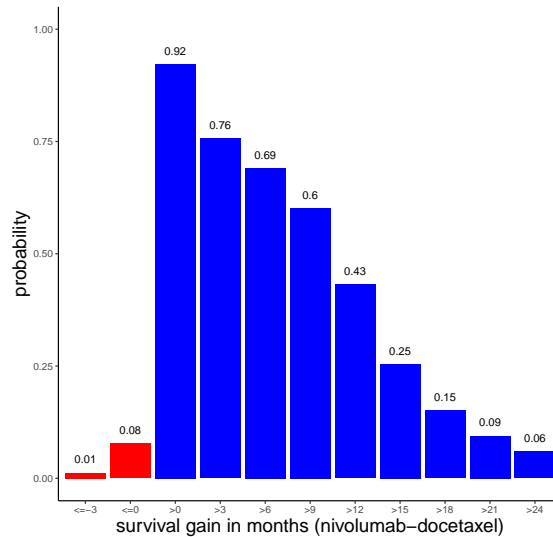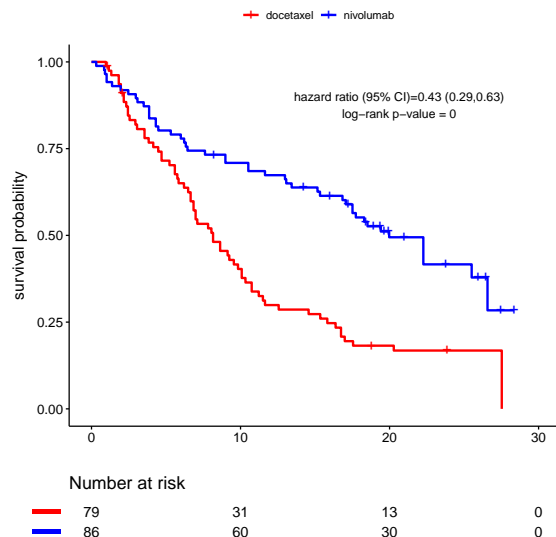
```
.. null device
..            1
```

```
pdf("appfig21gain.pdf",onefile=FALSE)
fig21.gain
dev.off()
```

```
.. null device
..            1
```

Figure 3: Kaplan-Meier curves and probabilities for survival gain for biomarker-negative (top) and biomarker-positive patients.

docetaxel nivolumab

survival probability

hazard ratio (95% CI)=0.43 (0.29,0.63)
log−rank p−value = 0

1.00

0.75

0.50

0.25

0.00

0   10   20   30

Number at risk

79   31   13   0
86   60   30   0

probability

0.92
0.76
0.69
0.6
0.43
0.25
0.15
0.09
0.06
0.01   0.08

1.00
0.75
0.50
0.25
0.00

<=−3 <=0 >0 >3 >6 >9 >12 >15 >18 >21 >24
survival gain in months (nivolumab−docetaxel)

sessionInfo()

```
.. R version 3.6.0 (2019-04-26)
.. Platform: x86_64-w64-mingw32/x64 (64-bit)
.. Running under: Windows 10 x64 (build 18363)
..
.. Matrix products: default
..
.. locale:
.. [1] LC_COLLATE=German_Switzerland.1252  LC_CTYPE=German_Switzerland.1252
.. [3] LC_MONETARY=German_Switzerland.1252 LC_NUMERIC=C
.. [5] LC_TIME=German_Switzerland.1252
..
.. attached base packages:
.. [1] stats      graphics  grDevices utils      datasets  methods    base
..
.. other attached packages:
.. [1] survminer_0.4.9    ggpubr_0.4.0       mvtnorm_1.0-10
.. [4] forcats_0.5.0      stringr_1.4.0      dplyr_1.0.5
.. [7] purrr_0.3.4        readr_1.3.1        tidyr_1.1.3
.. [10] tibble_3.0.1       tidyverse_1.3.0    survival_3.2-3
.. [13] ggplot2_3.3.3      clinpredict_0.1.2
..
.. loaded via a namespace (and not attached):
.. [1] httr_1.4.1         jsonlite_1.6       splines_3.6.0
.. [4] carData_3.0-4      modelr_0.1.8       assertthat_0.2.1
.. [7] highr_0.8          cellranger_1.1.0   yaml_2.2.0
.. [10] pillar_1.4.4       backports_1.1.4    lattice_0.20-41
.. [13] glue_1.4.2         digest_0.6.19      gridtext_0.1.4
.. [16] ggsignif_0.6.1     rvest_1.0.0        colorspace_1.4-1
.. [19] htmltools_0.4.0    Matrix_1.2-17      pkgconfig_2.0.2
.. [22] broom_0.7.6        haven_2.2.0        xtable_1.8-4
.. [25] scales_1.0.0       km.ci_0.5-2        openxlsx_4.2.3
.. [28] rio_0.5.26         KMsurv_0.1-5       generics_0.0.2
.. [31] car_3.0-10         ellipsis_0.3.0     withr_2.4.1
```

```
.. [34] cli_1.1.0           magrittr_1.5        crayon_1.3.4
.. [37] readxl_1.3.1        ggtext_0.1.1        evaluate_0.14
.. [40] fs_1.3.2            rstatix_0.7.0       xml2_1.3.2
.. [43] foreign_0.8-71      tools_3.6.0         data.table_1.12.2
.. [46] hms_0.5.3           lifecycle_1.0.0     munsell_0.5.0
.. [49] reprex_1.0.0        zip_2.1.1           compiler_3.6.0
.. [52] rlang_0.4.10        grid_3.6.0          rstudioapi_0.10
.. [55] labeling_0.3        rmarkdown_2.1       gtable_0.3.0
.. [58] abind_1.4-5         DBI_1.1.0           curl_4.3
.. [61] markdown_1.1        R6_2.4.0            gridExtra_2.3
.. [64] zoo_1.8-8           lubridate_1.7.10    knitr_1.24
.. [67] survMisc_0.5.5      stringi_1.4.3       Rcpp_1.0.6
.. [70] vctrs_0.3.6         dbplyr_2.1.0        tidyselect_1.1.0
.. [73] xfun_0.22
```

References

Borghaei, H., L. Paz-Ares, L. Horn, and others. 2015. "Nivolumab Versus Docetaxel in Advanced Nonsquamous Non–small-Cell Lung Cancer." *The New England Journal of Medicine* 373 (17): 1627–39.

Doksum, K. 1974. "Empirical probability plots and statistical inference for nonlinear models in the two-sample case." *Annals of Statistics* 2 (2): 267–77.

Lehmann, E.L. 1974. *Nonparameterics: Statistical Methods Based on Ranks*. Holden Day, San Francisco.