

1 A tutorial on Bayesian and Frequentist Event History Analyses for psychological
2 time-to-event data

3 Sven Panis¹ & Richard Ramsey¹

4 ¹ ETH Zürich

5 Author Note

6 Neural Control of Movement lab, Department of Health Sciences and Technology

7 (D-HEST).

8 The authors made the following contributions. Sven Panis: Conceptualization,
9 Writing - Original Draft Preparation, Writing - Review & Editing; Richard Ramsey:
10 Conceptualization, Writing - Review & Editing, Supervision.

11 Correspondence concerning this article should be addressed to Sven Panis, ETH
12 GLC, room G16.2, Gloriastrasse 37/39, 8006 Zürich. E-mail: sven.panis@hest.ethz.ch

13

Abstract

14 Time-to-event data such as response times, saccade latencies, and fixation durations are
15 ubiquitous in experimental psychology. The orthodox method for analysing such data –
16 comparing means with analysis-of-variance – is actually hiding a lot of information about
17 psychological effects, such as their onset time and duration, and whether they are
18 time-locked to stimuli. Such information can change key conclusions about psychological
19 processes and can be revealed by using distributional measures.

20 Here we provide a set of tutorials on how to implement one particular distributional
21 method known as discrete-time event history analysis, a.k.a. hazard analysis, duration
22 analysis, failure-time analysis, survival analysis, and transition analysis. We illustrate how
23 one can calculate the descriptive statistics, and how one can implement Bayesian and
24 frequentist regression models, using the R packages tidyverse, brms, and lme4. The R code
25 is publicly available on Github and OSF, and can easily be adapted for other data sets. We
26 discuss possible link functions and prior distributions, how to manage inter-individual
27 differences, implications for experimental design, and how to select among various options
28 when analysing time-to-event data using discrete-time hazard analysis.

29 *Keywords:* response times, event history analysis, Bayesian regression models

30 Word count: X

31 A tutorial on Bayesian and Frequentist Event History Analyses for psychological
 32 time-to-event data

33 **Introduction**

34 **Means versus distributional shapes**

35 In experimental psychology, it is still standard practice to analyse response times
 36 (RTs), saccade latencies, and fixation durations by calculating average performance across
 37 a series of trials. However, differences in means conceal when an experimental effect starts,
 38 how long it lasts, how it evolves over increasing waiting time, and whether its onset is
 39 time-locked to other events. Such information is useful not only for interpretation, but also
 40 for cognitive psychophysiology and computational model selection (Panis, Schmidt,
 41 Wolkersdorfer, & Schmidt, 2020). As a simple illustration, Figure 1 shows three examples
 42 of how an observed difference in mean response times (RTs) between two experimental
 43 conditions conceals differences in the shapes of the underlying RT distributions.

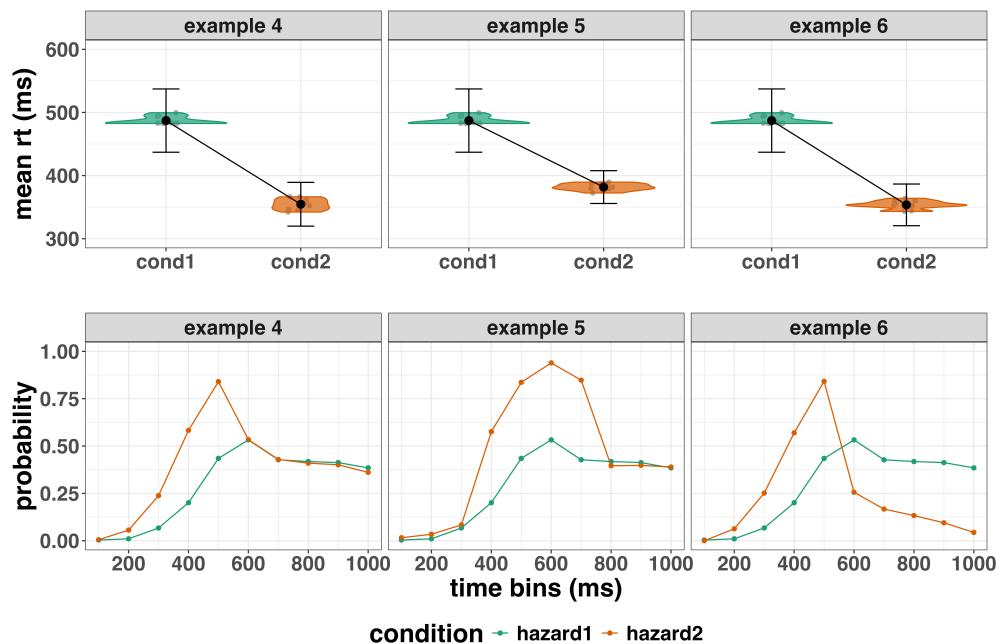


Figure 1. Means versus distributional shapes.

⁴⁴ In each example, the mean RT is lower in condition 2 compared to condition 1.
⁴⁵ However, the distributions in the first example show that the effect starts around 200 ms
⁴⁶ and is gone by 600 ms. In the second example, the distributional effect starts around 400
⁴⁷ ms and is gone by 800 ms. And in the third example, the distributional effect reverses
⁴⁸ around 550 ms.

⁴⁹ **Outline of the paper**

⁵⁰ In this paper we focus on a distributional method known as discrete-time event
⁵¹ history analysis, a.k.a. hazard analysis, duration analysis, failure-time analysis, survival
⁵² analysis, and transition analysis. We first provide a brief overview of hazard analysis to
⁵³ orient the reader to the basic concepts and ideas that we will use throughout the paper.
⁵⁴ However, this will remain relatively short, and for detailed treatment, see Singer and
⁵⁵ Willett (2003), Allison (1982), and Allison (2010).

⁵⁶ We then provide four different tutorials, each of which is written in R code and
⁵⁷ publicly available on Github and the Open Science Framework (OSF). The tutorials
⁵⁸ provide hands-on, concrete examples of key parts of the analytical process, so that others
⁵⁹ can apply the analyses to their own time-to-event data sets. In Tutorial 1 we illustrate how
⁶⁰ to calculate the descriptive statistics for a published data set when there is one
⁶¹ independent variable. The descriptive statistics are plotted, and we comment on their
⁶² interpretation. In Tutorial 2 we illustrate how one can fit Bayesian hazard models to the
⁶³ data. After selecting the best of four models, we plot the model-based effects and the
⁶⁴ model fits for a few subjects. In Tutorial 3 we illustrate how to fit hazard models in a
⁶⁵ frequentist framework. We compare the model-based effects between Bayesian and
⁶⁶ frequentist approaches. In Tutorial 4 we illustrate how to calculate the descriptive
⁶⁷ statistics when there are two independent variables.

68 Overview of hazard analysis

69 To apply event history analysis (EHA), one must be able to define the event of
70 interest (any qualitative change that can be situated in time, e.g., a button press, a saccade
71 onset, a fixation offset, etc.), time point zero (e.g., target stimulus onset, fixation onset),
72 and measure the passage of time between time point zero and event occurrence in discrete
73 or continuous time units. Both the definition of hazard and the type of models employed
74 depend on whether one is using continuous or discrete time units.

75 The shape of a distribution of waiting times can be described in multiple ways (Luce,
76 1991). Let RT be a continuous random variable denoting a particular person's response time
77 in a particular experimental condition. Because waiting times can only increase,
78 continuous-time EHA does not focus on the cumulative distribution function $F(t) = P(RT \leq t)$ and its derivative, the probability density function $f(t) = F(t)'$, but on the survivor
79 function $S(t) = P(RT > t)$ and the hazard rate function $\lambda(t) = f(t)/S(t)$. The hazard rate
80 function gives you the instantaneous rate of event occurrence at time point t , given that
81 the event has not occurred yet.
82

83 Similarly, after dividing time in discrete, contiguous time bins indexed by t , let RT be
84 a discrete random variable denoting the rank of the time bin in which a particular person's
85 response occurs in a particular experimental condition. Discrete-time EHA focuses on the
86 discrete-time hazard function $h(t) = P(RT = t | RT \geq t)$ and the discrete-time survivor
87 function $S(t) = P(RT > t) = [1-h(t)].[1-h(t-1)].[1-h(t-2)] \dots [1-h(1)]$, and not on the
88 probability mass function $p(t) = h(t).S(t-1)$ and the cumulative distribution function $F(t) = 1-S(t)$. The discrete-time hazard probability function gives you the probability that the
89 event occurs (sometime) in bin t , given that the event has not occurred yet in previous
90 bins. Unlike the discrete-time hazard function, which assesses the unique risk associated
91 with each time bin, the discrete-time survivor function cumulates the bin-by-bin risks of
92 event *nonoccurrence*.

94 For two-choice RT data, the discrete-time hazard function can be extended with the
95 discrete-time conditional accuracy function $ca(t) = P(\text{correct} \mid RT = t)$, which gives you
96 the probability that a response is correct given that it has been emitted in time bin t
97 (Allison, 2010; Kantowitz & Pachella, 2021; Wickelgren, 1977). This latter function is also
98 known as the micro-level speed-accuracy tradeoff function.

99 Statisticians and mathematical psychologists recommend focusing on the hazard
100 function when analyzing time-to-event data for various reasons. First, as discussed by
101 Holden, Van Orden, and Turvey (2009), “probability density functions can appear nearly
102 identical, both statistically and to the naked eye, and yet are clearly different on the basis
103 of their hazard functions (but not vice versa). Hazard functions are thus more diagnostic
104 than density functions” (p. 331). Second, because RT distributions may differ from one
105 another in multiple ways, Townsend (1990) developed a dominance hierarchy of statistical
106 differences between two arbitrary distributions A and B. For example, if $F_A(t) > F_B(t)$ for
107 all t , then both cumulative distribution functions are said to show a complete ordering.
108 Townsend (1990) showed that a complete ordering on the hazard functions $-\lambda_A(t) > \lambda_B(t)$
109 for all t —implies a complete ordering on both the cumulative distribution and survivor
110 functions $-F_A(t) > F_B(t)$ and $S_A(t) < S_B(t)$ —which in turn implies an ordering on the
111 mean latencies—mean A < mean B. In contrast, an ordering on two means does not imply
112 a complete ordering on the corresponding $F(t)$ and $S(t)$ functions, and a complete ordering
113 on these latter functions does not imply a complete ordering on the corresponding hazard
114 functions. This means that stronger conclusions can be drawn from data when comparing
115 the hazard functions using EHA. For example, when mean A < mean B, the hazard
116 functions might show a complete ordering (i.e., for all t), a partial ordering (e.g., only for t
117 > 300 ms, or only for $t < 500$ ms), or they may cross each other one or more times. Third,
118 EHA does not discard right-censored observations when estimating hazard functions, that
119 is, trials for which we do not observe a response during the data collection period so that
120 we only know that the RT must be larger than some value. This is important because

121 although a few right-censored observations are inevitable in most RT tasks, a lot of
122 right-censored observations are expected in experiments on masking, the attentional blink,
123 and so forth. Fourth, hazard modeling allows incorporating time-varying explanatory
124 covariates such as heart rate, electroencephalogram (EEG) signal amplitude, gaze location,
125 etc. (Allison, 2010) which is useful for cognitive psychophysiology (Meyer, Osman, Irwin, &
126 Yantis, 1988). Finally, as explained by Kelso, Dumas, and Tognoli (2013), it is crucial to
127 first have a precise description of the macroscopic behavior of a system (here: $h(t)$ and
128 $ca(t)$ functions) in order to know what to derive on the microscopic level. For example,
129 fitting parametric functions or computational models to data without studying the shape
130 of the discrete-time $h(t)$ and $ca(t)$ functions can miss important features in the data
131 (Panis, Moran, Wolkersdorfer, & Schmidt, 2020; Panis & Schmidt, 2016).

132 We focus on factorial within-subject designs in which a large number of observations
133 are made on a relatively small number of participants (small- N designs). This approach
134 emphasizes the precision and reproducibility of data patterns at the individual participant
135 level to increase the inferential validity of the design (Baker et al., 2021; Smith & Little,
136 2018). In contrast to the large- N design that averages across many participants without
137 being able to scrutinize individual data patterns, small- N designs retain crucial
138 information about the data patterns of individual observers. This is of great advantage
139 whenever participants differ systematically in their strategies or in the time-courses of their
140 effects, so that blindly averaging them would lead to misleading data patterns. Indeed,
141 Smith and Little (2018) argue that, “if psychology is to be a mature quantitative science,
142 then its primary theoretical aim should be to investigate systematic functional
143 relationships as they are manifested at the individual participant level” (p. 2083). Note
144 that because statistical power derives both from the number of participants and from the
145 number of repeated measures per participant and condition, small- N designs can have
146 excellent power (Baker et al., 2021; Smith & Little, 2018).

We used R (Version 4.4.0; R Core Team, 2024)¹ for all reported analyses. Web links printed in bold. The content of the tutorials is mainly based on Allison (2010), Singer Willett (2003), McElreath (2018), Kurz (2023a), and Kurz (2023b).

Tutorial 1: Calculating descriptive statistics using a life table

To illustrate how to quickly set up life tables for calculating the descriptive statistics of discrete time), we use a published data set on masked response priming from Panis and Schmidt (2016), available on **ResearchGate**. In their first experiment, Panis and Schmidt (2016) presented a double arrow for 94 ms that pointed left or right as the target stimulus with an onset at time point zero in each trial. Participants had to indicate the direction in which the double arrow pointed using their corresponding index finger, within 800 ms after target onset. Response time and accuracy were recorded on each trial. Prime type (blank, congruent, incongruent) and mask type were manipulated. Here we focus on the subset of trials in which no mask was presented. The 13-ms prime stimulus

160 was a double arrow with onset at -187 ms for the congruent (same direction as target) and
 161 incongruent (opposite direction as target) prime conditions.

162 After loading in the data file, one has to (a) supply required column names, and (b)
 163 specify the factor condition with the correct levels and labels. The required column names
 164 are as follows:

- 165 • “pid”, indicating unique participant IDs;
- 166 • “trial”, indicating each unique trial per participant;
- 167 • “condition”, a factor indicating the levels of the independent variable (1, 2, ...) and
 168 the corresponding labels;
- 169 • “rt”, indicating the response times in ms;
- 170 • “acc”, indicating the accuracies (1/0).

171 In the code of Tutorial 1, this is accomplished as follows.

```
data_wr <- read_csv("../Tutorial_1_descriptive_stats/data/DataExp1_6subjects_wrangled.csv")
colnames(data_wr) <- c("pid", "bl", "tr", "condition", "resp", "acc", "rt", "trial")
data_wr <- data_wr %>%
  mutate(condition = condition + 1, # original levels were 0, 1, 2.
        condition = factor(condition, levels=c(1,2,3), labels=c("blank", "congruent", "incongruent")))
```

172 To set up the life tables and plots of the discrete-time functions $h(t)$, $S(t)$ and $ca(t)$
 173 using functional programming, one has to nest the data within participants using the
 174 `group_nest()` function, and supply a user-defined censoring time and bin width to our
 175 function “`censor()`”, as follows.

```
data_nested <- data_wr %>% group_nest(pid)
data_final <- data_nested %>%
  mutate(censored = map(data, censor, 600, 40)) %>% # ! user input: censoring time, and bin width
  mutate(ptb_data = map(censored, ptb)) %>% # create person-trial-bin dataset
  mutate(lifetable = map(ptb_data, setup_lt)) %>% # create life tables without ca(t)
  mutate(condacc = map(censored, calc_ca)) %>% # calculate ca(t)
  mutate(lifetable_ca = map2(lifetable, condacc, join_lt_ca)) %>% # create life tables with ca(t)
  mutate(plot = map2(.x = lifetable_ca, .y = pid, plot_eha, 1)) # create plots
```

176 Note that the censoring time should be a multiple of the bin width (both in ms). The

177 censoring time should be a time point after which no informative responses are expected

178 anymore. In experiments that implement a response deadline in each trial the censoring

179 time can equal that deadline time point. Trials with a RT larger than the censoring time,

180 or trials in which no response is emitted during the data collection period, are treated as

181 right-censored observations in EHA. In other words, these trials are not discarded, because

182 they contain the information that the event did not occur before the censoring time.

183 Removing such trials before calculating the mean event time can introduce a sampling bias.

184 The person-trial-bin oriented dataset has one row for each time bin of each trial that is at

185 risk for event occurrence. The variable “event” in the person-trial-bin oriented data set

186 indicates whether a response occurs (1) or not (0) for each bin. When creating the plots

187 using our function `plot_eha()`, some warning messages will likely be generated, like these:

- Removed 2 rows containing missing values or values outside the scale range

189 (`geom_line()`).

- Removed 2 rows containing missing values or values outside the scale range

191 (`geom_point()`).

- Removed 2 rows containing missing values or values outside the scale range

193 (`geom_segment()`).

194 The warning messages are generated because some bins have no hazard and $ca(t)$

195 estimates, and no error bars. They can thus safely be ignored. One can now inspect

196 different aspects, including the life table for a particular condition of a particular subject,

197 and a plot of the different functions for a particular participant.

198 Table 1 shows the life table for condition “blank” (no prime stimulus presented) -

199 compare to Figure 1. A life table includes for each time bin, the risk set (number of trials

200 that are event-free at the start of the bin), the number of observed events, and the

201 estimates of $h(t)$, $S(t)$, $ca(t)$ and their estimated standard errors (se). At time point zero,
202 no events can occur and therefore $h(t)$ and $ca(t)$ are undefined.

203 Figure 1 displays the discrete-time hazard, survivor, and conditional accuracy
204 functions for each prime condition for participant 6. By using discrete-time $h(t)$ functions
205 of event occurrence - in combination with $ca(t)$ functions for two-choice tasks - one can
206 provide an unbiased, time-varying, and probabilistic description of the latency and
207 accuracy of responses based on all trials of any data set.

208 For example, for participant 6, the estimated hazard values in bin (240,280] are 0.03,
209 0.17, and 0.20 for the blank, congruent, and incongruent prime conditions, respectively. In
210 other words, when the waiting time has increased until *240 ms* after target onset, then the
211 conditional probability of response occurrence in the next 40 ms is more than five times
212 larger for both prime-present conditions, compared to the blank prime condition.

213 Furthermore, the estimated conditional accuracy values in bin (240,280] are 0.29, 1,
214 and 0 for the blank, congruent, and incongruent prime conditions, respectively. In other
215 words, if a response is emitted in bin (240,280], then the probability that it is correct is
216 estimated to be 0.29, 1, and 0 for the blank, congruent, and incongruent prime conditions,
217 respectively.

Table 1

The life table for the blank prime condition of participant 6.

bin	risk_set	events	hazard	se_haz	survival	se_surv	ca	se_ca
0.00	220.00	NA	NA	NA	1.00	0.00	NA	NA
40.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
80.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
120.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
160.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
200.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
240.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
280.00	220.00	7.00	0.03	0.01	0.97	0.01	0.29	0.17
320.00	213.00	13.00	0.06	0.02	0.91	0.02	0.77	0.12
360.00	200.00	26.00	0.13	0.02	0.79	0.03	0.92	0.05
400.00	174.00	40.00	0.23	0.03	0.61	0.03	1.00	0.00
440.00	134.00	48.00	0.36	0.04	0.39	0.03	0.98	0.02
480.00	86.00	37.00	0.43	0.05	0.22	0.03	1.00	0.00
520.00	49.00	32.00	0.65	0.07	0.08	0.02	1.00	0.00
560.00	17.00	9.00	0.53	0.12	0.04	0.01	1.00	0.00
600.00	8.00	4.00	0.50	0.18	0.02	0.01	1.00	0.00

Note. The column named “bin” indicates the endpoint of each time bin (in ms), and includes time point zero. For example the first bin is (0,40] with the starting point excluded and the endpoint included. se = standard error. ca = conditional accuracy.

Subject 6

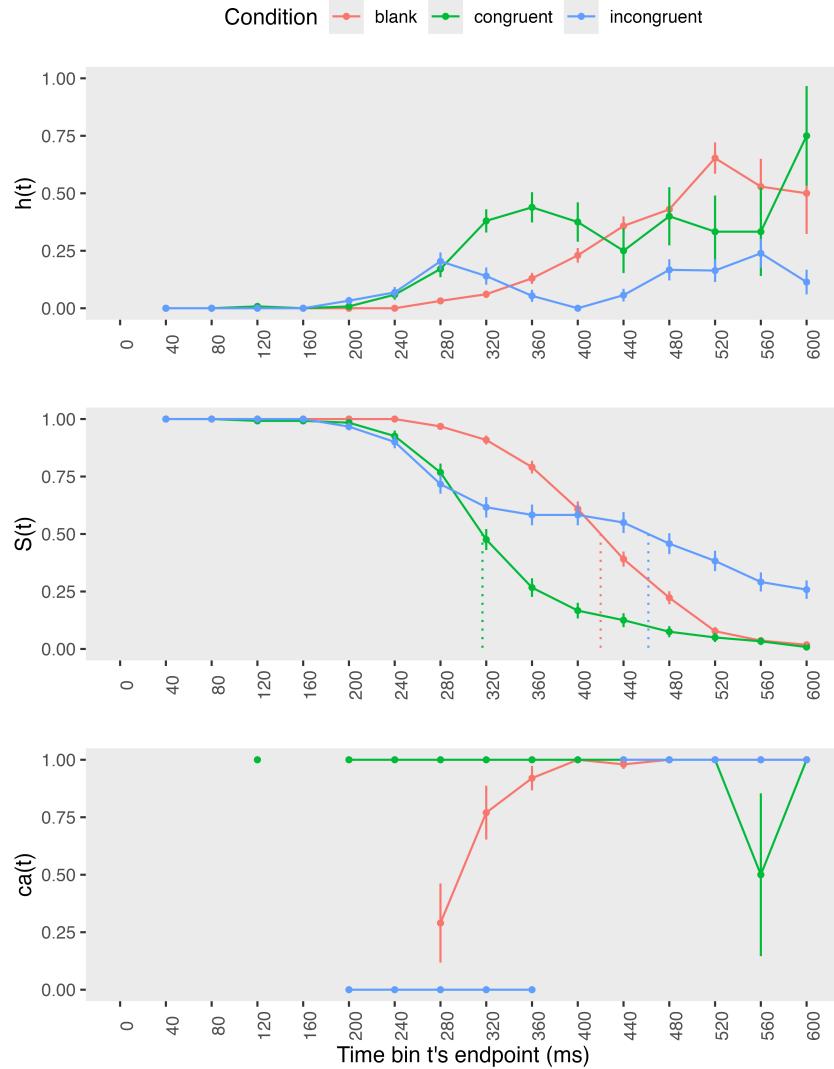


Figure 2. Estimated discrete-time hazard, survivor, and conditional accuracy functions for participant 6, as a function of the passage of discrete waiting time.

218 However, when the waiting time has increased until 400 ms after target onset, then
 219 the conditional probability of response occurrence in the next 40 ms is estimated to be
 220 0.36, 0.25, and 0.06 for the blank, congruent, and incongruent prime conditions,
 221 respectively. And when a response does occur in bin (400,440], then the probability that it
 222 is correct is estimated to be 0.98, 1, and 1 for the blank, congruent, and incongruent prime

223 conditions, respectively.

224 These results show that this participant is initially responding to the prime even
225 though (s)he was instructed to only respond to the target, that response competition
226 emerges in the incongruent prime condition around 300 ms, and that only later response
227 are fully controlled by the target stimulus. Qualitatively similar results were obtained for
228 the other five participants. Also, in their second Experiment, Panis and Schmidt (2016)
229 showed that the negative compatibility effect in the mask-present conditions is time-locked
230 to mask onset. This example shows that a simple difference between two means fails to
231 reveal the dynamic behavior people display in many experimental paradigms (Panis, 2020;
232 Panis, Moran, et al., 2020; Panis & Schmidt, 2022; Panis, Torfs, Gillebert, Wagemans, &
233 Humphreys, 2017; Panis & Wagemans, 2009; Schmidt, Panis, Wolkersdorfer, & Vorberg,
234 2022). In other words, statistically controlling for the passage of time during data analysis
235 is equally important as experimental control during the design of an experiment, to better
236 understand human behavior in experimental paradigms. As we will show in Tutorials 2 and
237 3, statistical models for $h(t)$ can be implemented as generalized linear mixed regression
238 models predicting event occurrence (1/0) in each bin of a selected time range.

239 **Tutorial 2: Fitting Bayesian hazard models**

240 When you want to study how hazard depends on various predictors, you can fit
241 regression models to the data (Singer & Willett, 2003). There are two analytic decisions
242 one has to make. First, one has to select an analysis time range, i.e., a contiguous set of
243 bins for which there is enough data for each participant. Second, one can choose the logit
244 link function which transforms a (hazard) probability into the log of the odds ratio, or the
245 complementary log-log (cloglog) link function, which yields the logarithm of the negated
246 logarithm of the probability of event *nonoccurrence*. An important difference between
247 these two link functions is that cloglog provides a discrete-time hazard model that has a
248 built-in proportional hazards assumption, while logit provides a proportional odds

assumption (see below). The cloglog link is preferred over the logit link when events can occur in principle at any time point within a bin, which is the case for RT data (Singer & Willett, 2003). Third, one has to choose a specification of the effect of discrete TIME (i.e., the time bin index t). One can choose a general specification (one intercept per bin) or a functional specification, such as a polynomial one.

An example discrete-time hazard model with three predictors (TIME, X1, X2) and the cloglog link function can be written as follows:

$$\text{cloglog}[h(t)] = \ln(-\ln[1-h(t)]) = [\alpha_1 \text{ONE} + \alpha_2(\text{TIME}-1) + \alpha_3(\text{TIME}-1)^2] + [\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_2(\text{TIME}-1)].$$

The main predictor variable TIME is the time bin index t that is centered on value 1 in this example. The first set of terms within brackets, the alpha parameters multiplied by their polynomial specifications of (centered) time, represents the shape of the baseline cloglog-hazard function (i.e., when all predictors X_i take on a value of zero). The second set of terms (the beta parameters) represents the vertical shift in the baseline cloglog-hazard for a 1 unit increase in the respective predictor. Predictors can be discrete, continuous, and time-varying or time-invariant. For example, the effect of a 1 unit increase in X_1 is to vertically shift the whole baseline cloglog-hazard function by β_1 cloglog-hazard units. However, if the predictor interacts linearly with time (see X_2 in the example), then the effect of a 1 unit increase in X_2 is to vertically shift the predicted cloglog-hazard in bin 1 by β_2 cloglog-hazard units (when $\text{TIME}-1 = 0$), in bin 2 by $\beta_2 + \beta_3$ cloglog-hazard units (when $\text{TIME}-1 = 1$), and so forth. To interpret the effects of the predictors, the parameter estimates are exponentiated, resulting in a hazard ratio (due to the use of the cloglog link).

In the case of a large- N design without repeated measurements, the parameters of a discrete-time hazard model can be estimated using standard logistic regression software (after expanding the typical person-trial-oriented data set into a person-trial-bin-oriented data set (see Tutorial 1); Allison (2010)). When there is clustering in the data, as in the

275 case of a small- N design with repeated measurements, the parameters of a discrete-time
 276 hazard model can be estimated using population-averaged methods (e.g., Generalized
 277 Estimating Equations), and Bayesian or frequentist generalized linear mixed models
 278 (Allison, 2010).

279 In this second tutorial we illustrate how to fit a Bayesian hazard regression model for
 280 the masked response priming data set used in the first tutorial. In general, there are three
 281 assumptions one can make or relax when adding experimental predictor variables: The
 282 linearity assumption for continuous predictors (the effect of a 1 unit change is the same
 283 anywhere on the scale), the additivity assumption (predictors do not interact), and the
 284 proportionality assumption (predictors do not interact with TIME).

285 First, we select the analysis range (200,600] and the cloglog link, and use a
 286 polynomial to specify the effect of TIME in the “blank” prime condition. Second, based on
 287 previous work (Panis, 2020; Panis, Moran, et al., 2020; Panis & Schmidt, 2022; Panis et al.,
 288 2017; Panis & Wagemans, 2009) and because cognition is likely the behavior of a non-linear
 289 dynamical system [ref], we relax all three assumptions, as follows:

```
# load data
ptb_data <- read_csv("../Tutorial_1_descriptive_stats/data/inputfile_hazard_modeling.csv")
# select analysis time range: (200,600] with 10 bins (time bin ranks 6 to 15)
ptb_data <- ptb_data %>% filter(period > 5)
# create factor condition, with "blank" as the reference level
ptb_data <- ptb_data %>% mutate(condition = factor(condition, labels = c("blank", "congruent","incongruent")))
# center discrete TIME (period) on bin 9, and trial on trial 1000
ptb_data <- ptb_data %>% mutate(period_9 = period - 9,
                                    trial_c = (trial - 1000)/1000)
# remove unnecessary columns before fitting a model
ptb_data <- ptb_data %>% select(-c(bl,tr,trial,period)) # 12840 obs. of 5 variables

priors <- c(
  set_prior("normal(0, 1)", class = "b"), # for beta parameters
  set_prior("student_t(7.61, 0, 1.57)", class = "b", coef = "Intercept"), # flat prior for intercept on hazard scale
  set_prior("normal(0, 1)", class = "sd"), # for standard deviation of RE
  set_prior("lkj(2)", class = "cor") # for correlations between RE
)
```

```
#plan(multicore)
#model_full_RE <-
#  brm(data = ptb_data,
#       family = binomial(link="cloglog"),
#       event ~ trials(1) ~ 0 + Intercept +
#                           condition*period_9*trial_c +
#                           condition*I(period_9^2) +
#                           condition*I(period_9^3) +
#                           (1 + condition*period_9*trial_c +
#                           condition*I(period_9^2) +
#                           condition*I(period_9^3) | pid),
#       prior = priors,
#       chains = 4, cores = 4, iter = 3000, warmup = 1000,
#       control = list(adapt_delta = 0.999, step_size = 0.04, max_treedepth = 12),
#       seed = 12, init = "O",
#       file = "../Tutorial_2_Bayesian/models/model_full_RE")
```

290 To test whether (centered) trial number affects behavior, we fit a model without the
 291 variable trial_c.

292 Use WAIC to compare models.

293 Plot the effects of congruent and incongruent for each time bin for the selected model.

294 Plot the model-based hazard and survivor functions.

295 **Tutorial 3: Fitting Frequentist hazard models**

296 In this third tutorial we illustrate how to fit a frequentist hazard regression model for
 297 the data set used in the first tutorial.

298 **Tutorial 4: Calculating descriptive statistics when there are two independent 299 variables**

300 In this final tutorial we illustrate how to calculate and plot the descriptive statistics
 301 for the full data set of Experiment 1 of Panis and Schmidt (2016), which includes two
 302 independent variables: mask type and prime type.

303

Discussion

304 We noticed that many researchers are still reluctant to abandon analysis-of-variance
305 and switch to event history analysis when analysing time-to-event data. By providing this
306 tutorial, we hope that researchers will start using hazard analysis more often. While we
307 focused on within-subject factorial small- N designs, event history analysis has been
308 developed to also encompass other designs.

309 **Advantages of hazard functions**

310 **Individual differences**

- 311 • role of response deadlines, low-level vs. higher-level processes, overlap between
- 312 distributions
- 313 • clustering algorithms based on $h(t)$ and $ca(t)$ data
- 314 •

315 **Cognitive psychophysiology and computational model selection**

316 **Power analysis**

- 317 • example repo on github

318 **Preregistration**

- 319 • example preregistration for knot data

320

Conclusions

321

References

- 322 Allison, P. D. (1982). Discrete-Time Methods for the Analysis of Event Histories.
323 *Sociological Methodology*, 13, 61. <https://doi.org/10.2307/270718>
- 324 Allison, P. D. (2010). *Survival analysis using SAS: A practical guide* (2. ed). Cary, NC:
325 SAS Press.
- 326 Aust, F. (2019). *Citr: 'RStudio' add-in to insert markdown citations*. Retrieved from
327 <https://github.com/crsh/citr>
- 328 Aust, F., & Barth, M. (2023). *papaja: Prepare reproducible APA journal articles with R*
329 *Markdown*. Retrieved from <https://github.com/crsh/papaja>
- 330 Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., &
331 Andrews, T. J. (2021). Power contours: Optimising sample size and precision in
332 experimental psychology and human neuroscience. *Psychological Methods*, 26(3),
333 295–314. <https://doi.org/10.1037/met0000337>
- 334 Barth, M. (2023). *tinylabes: Lightweight variable labels*. Retrieved from
335 <https://cran.r-project.org/package=tinylabes>
- 336 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects
337 models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
338 <https://doi.org/10.18637/jss.v067.i01>
- 339 Bates, D., Maechler, M., & Jagan, M. (2024). *Matrix: Sparse and dense matrix classes and*
340 *methods*. Retrieved from <https://CRAN.R-project.org/package=Matrix>
- 341 Bengtsson, H. (2021). A unifying framework for parallel and distributed processing in r
342 using futures. *The R Journal*, 13(2), 208–227. <https://doi.org/10.32614/RJ-2021-048>
- 343 Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan.
344 *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- 345 Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms.
346 *The R Journal*, 10(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>
- 347 Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal*

- 348 *of Statistical Software*, 100(5), 1–54. <https://doi.org/10.18637/jss.v100.i05>
- 349 Eddelbuettel, D., & Balamuta, J. J. (2018). Extending R with C++: A Brief Introduction
350 to Rcpp. *The American Statistician*, 72(1), 28–36.
351 <https://doi.org/10.1080/00031305.2017.1375990>
- 352 Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal
353 of Statistical Software*, 40(8), 1–18. <https://doi.org/10.18637/jss.v040.i08>
- 354 Gabry, J., Češnovar, R., Johnson, A., & Broder, S. (2024). *Cmdstanr: R interface to
355 'CmdStan'*. Retrieved from <https://github.com/stan-dev/cmdstanr>
- 356 Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization
357 in bayesian workflow. *J. R. Stat. Soc. A*, 182, 389–402.
358 <https://doi.org/10.1111/rssa.12378>
- 359 Girard, J. (2024). *Standist: What the package does (one line, title case)*. Retrieved from
360 <https://github.com/jmgirard/standist>
- 361 Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate.
362 *Journal of Statistical Software*, 40(3), 1–25. Retrieved from
363 <https://www.jstatsoft.org/v40/i03/>
- 364 Holden, J. G., Van Orden, G. C., & Turvey, M. T. (2009). Dispersion of response times
365 reveals cognitive dynamics. *Psychological Review*, 116(2), 318–342.
366 <https://doi.org/10.1037/a0014849>
- 367 Kantowitz, B. H., & Pachella, R. G. (2021). The Interpretation of Reaction Time in
368 Information-Processing Research 1. *Human Information Processing*, 41–82.
369 <https://doi.org/10.4324/9781003176688-2>
- 370 Kay, M. (2023). *tidybayes: Tidy data and geoms for Bayesian models*.
371 <https://doi.org/10.5281/zenodo.1308151>
- 372 Kelso, J. A. S., Dumas, G., & Tognoli, E. (2013). Outline of a general theory of behavior
373 and brain coordination. *Neural Networks: The Official Journal of the International
374 Neural Network Society*, 37, 120–131. <https://doi.org/10.1016/j.neunet.2012.09.003>

- 375 Kurz, A. S. (2023a). *Applied longitudinal data analysis in brms and the tidyverse* (version
376 0.0.3). Retrieved from <https://bookdown.org/content/4253/>
- 377 Kurz, A. S. (2023b). *Statistical rethinking with brms, ggplot2, and the tidyverse: Second*
378 *edition* (version 0.4.0). Retrieved from <https://bookdown.org/content/4857/>
- 379 Luce, R. D. (1991). *Response times: Their role in inferring elementary mental organization*
380 (1. issued as paperback). Oxford: Univ. Press.
- 381 McElreath, R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and*
382 *Stan* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315372495>
- 383 Meyer, D. E., Osman, A. M., Irwin, D. E., & Yantis, S. (1988). Modern mental
384 chronometry. *Biological Psychology*, 26(1-3), 3–67.
385 [https://doi.org/10.1016/0301-0511\(88\)90013-0](https://doi.org/10.1016/0301-0511(88)90013-0)
- 386 Müller, K., & Wickham, H. (2023). *Tibble: Simple data frames*. Retrieved from
387 <https://CRAN.R-project.org/package=tibble>
- 388 Neuwirth, E. (2022). *RColorBrewer: ColorBrewer palettes*. Retrieved from
389 <https://CRAN.R-project.org/package=RColorBrewer>
- 390 Panis, S. (2020). How can we learn what attention is? Response gating via multiple direct
391 routes kept in check by inhibitory control processes. *Open Psychology*, 2(1), 238–279.
392 <https://doi.org/10.1515/psych-2020-0107>
- 393 Panis, S., Moran, R., Wolkersdorfer, M. P., & Schmidt, T. (2020). Studying the dynamics
394 of visual search behavior using RT hazard and micro-level speed–accuracy tradeoff
395 functions: A role for recurrent object recognition and cognitive control processes.
396 *Attention, Perception, & Psychophysics*, 82(2), 689–714.
397 <https://doi.org/10.3758/s13414-019-01897-z>
- 398 Panis, S., Schmidt, F., Wolkersdorfer, M. P., & Schmidt, T. (2020). Analyzing Response
399 Times and Other Types of Time-to-Event Data Using Event History Analysis: A Tool
400 for Mental Chronometry and Cognitive Psychophysiology. *I-Perception*, 11(6),
401 2041669520978673. <https://doi.org/10.1177/2041669520978673>

- 402 Panis, S., & Schmidt, T. (2016). What Is Shaping RT and Accuracy Distributions? Active
403 and Selective Response Inhibition Causes the Negative Compatibility Effect. *Journal of*
404 *Cognitive Neuroscience*, 28(11), 1651–1671. https://doi.org/10.1162/jocn_a_00998
- 405 Panis, S., & Schmidt, T. (2022). When does “inhibition of return” occur in spatial cueing
406 tasks? Temporally disentangling multiple cue-triggered effects using response history
407 and conditional accuracy analyses. *Open Psychology*, 4(1), 84–114.
408 <https://doi.org/10.1515/psych-2022-0005>
- 409 Panis, S., Torfs, K., Gillebert, C. R., Wagemans, J., & Humphreys, G. W. (2017).
410 Neuropsychological evidence for the temporal dynamics of category-specific naming.
411 *Visual Cognition*, 25(1-3), 79–99. <https://doi.org/10.1080/13506285.2017.1330790>
- 412 Panis, S., & Wagemans, J. (2009). Time-course contingencies in perceptual organization
413 and identification of fragmented object outlines. *Journal of Experimental Psychology:*
414 *Human Perception and Performance*, 35(3), 661–687.
415 <https://doi.org/10.1037/a0013547>
- 416 Pedersen, T. L. (2024). *Patchwork: The composer of plots*. Retrieved from
417 <https://CRAN.R-project.org/package=patchwork>
- 418 Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in s and s-PLUS*. New York:
419 Springer. <https://doi.org/10.1007/b98882>
- 420 R Core Team. (2024). *R: A language and environment for statistical computing*. Vienna,
421 Austria: R Foundation for Statistical Computing. Retrieved from
422 <https://www.R-project.org/>
- 423 Schmidt, T., Panis, S., Wolkersdorfer, M. P., & Vorberg, D. (2022). Response inhibition in
424 the Negative Compatibility Effect in the absence of inhibitory stimulus features. *Open*
425 *Psychology*, 4(1), 219–230. <https://doi.org/10.1515/psych-2022-0012>
- 426 Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling*
427 *Change and Event Occurrence*. Oxford, New York: Oxford University Press.
- 428 Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design.

- 429 *Psychonomic Bulletin & Review*, 25(6), 2083–2101.
- 430 <https://doi.org/10.3758/s13423-018-1451-8>
- 431 Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical
432 distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, 108(3),
433 551–567. <https://doi.org/10.1037/0033-2909.108.3.551>
- 434 Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics.
435 *Acta Psychologica*, 41(1), 67–85. [https://doi.org/10.1016/0001-6918\(77\)90012-9](https://doi.org/10.1016/0001-6918(77)90012-9)
- 436 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New
437 York. Retrieved from <https://ggplot2.tidyverse.org>
- 438 Wickham, H. (2023a). *Forcats: Tools for working with categorical variables (factors)*.
439 Retrieved from <https://forcats.tidyverse.org/>
- 440 Wickham, H. (2023b). *Stringr: Simple, consistent wrappers for common string operations*.
441 Retrieved from <https://stringr.tidyverse.org>
- 442 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ...
443 Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43),
444 1686. <https://doi.org/10.21105/joss.01686>
- 445 Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A
446 grammar of data manipulation*. Retrieved from <https://dplyr.tidyverse.org>
- 447 Wickham, H., & Henry, L. (2023). *Purrr: Functional programming tools*. Retrieved from
448 <https://purrr.tidyverse.org/>
- 449 Wickham, H., Hester, J., & Bryan, J. (2024). *Readr: Read rectangular text data*. Retrieved
450 from <https://readr.tidyverse.org>
- 451 Wickham, H., Vaughan, D., & Girlich, M. (2024). *Tidyr: Tidy messy data*. Retrieved from
452 <https://tidyr.tidyverse.org>