Permutation test for general dependent truncation

EcoSta2019 @Taichung, Taiwan

Sy Han (Steven) Chiou

Joint work with
Dr. Jing Qian, Dr. Elizabeth Mormino, and Dr. Rebecca Betensky.

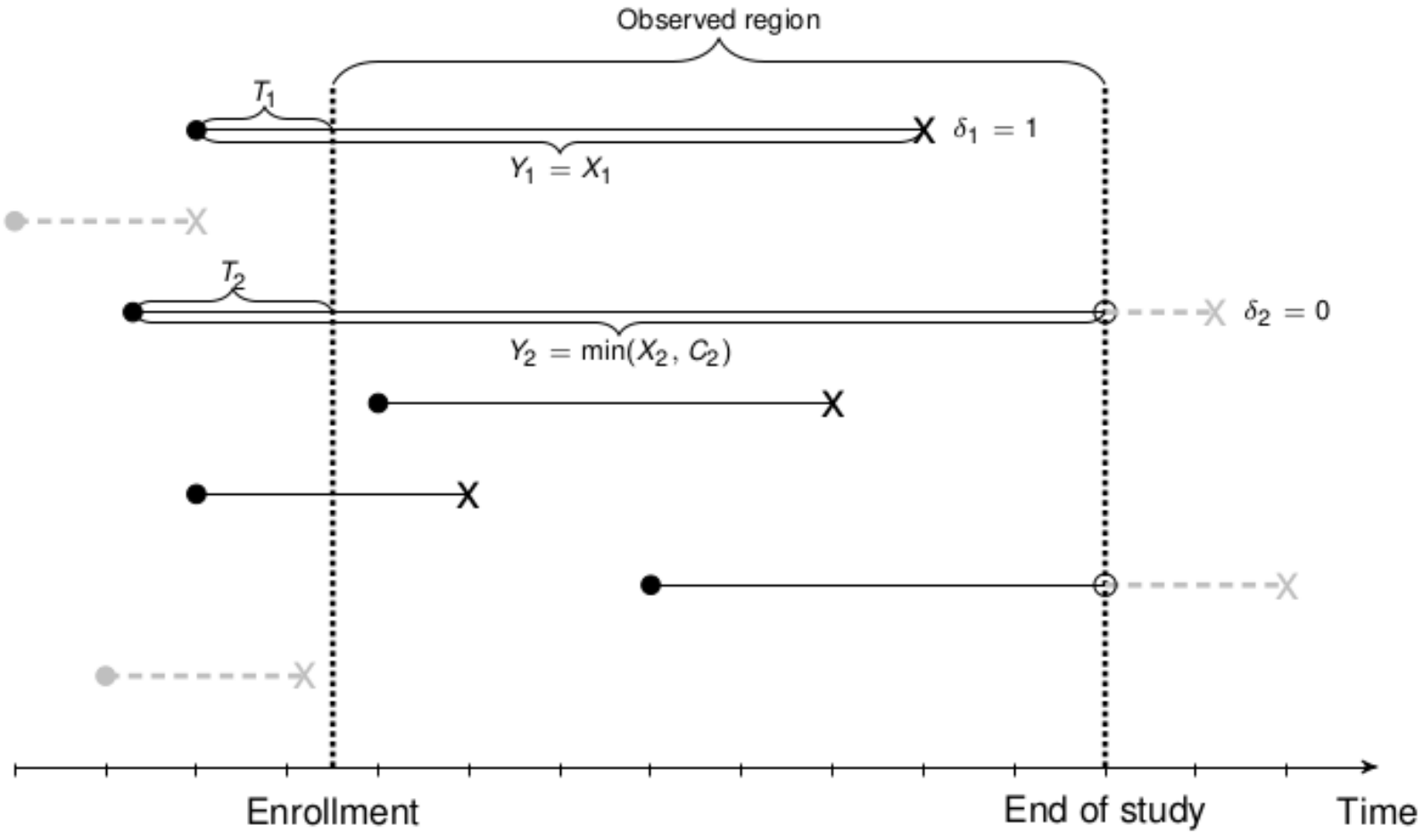
THE UNIVERSITY OF TEXAS AT DALLAS
School of Natural Sciences and Mathematics

Table of contents

Background

Left-truncated right-censored survival data

- An example of a left-truncated right-censored survival data



Notations

- X is the failure or event time
- T is the truncation time for X
- C is the right censoring time
- Y is the observed failure time: $Y = \min(X, C)$
- δ is the censoring indicator: $\delta = 1$ if $X \leq C$ and 0 otherwise.
- The observed data are $(Y, T, \delta \mid Y \geq T)$

Statistical challenges

- No information is observed when $X \leq T$
- Independence in the observable region, i.e., quasi-independence.
- Substantial bias if falsely assumed.
- Possible dependence could be non-monotonic.

Quasi-independence

- Quasi-independence is independence in the observable region:

$$h(x, t|X > T) \propto f(x)g(t), x \geq t,$$

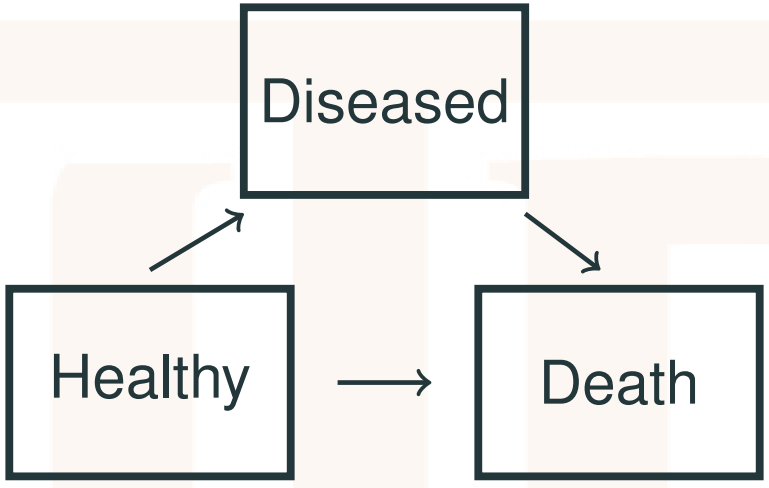
  where, $h(\cdot, \cdot)$, $f(\cdot)$, and $g(\cdot)$ are probability density* functions [Vakulenko-Lagun et al., 2018+].
- This condition can be tested, but available tests are limited.

Existing tests

- Tsai [1990]: conditional Kendall's tau (τ_c) based permutation test.
- Jones and Crowley [1992]: Cox model based tests
- Chen et al. [1996]: conditional Pearson correlation coefficient
- Martin and Betensky [2005]: general truncation via U-statistics
- Austin and Betensky [2014]: modifies τ_c via inverse probability weighting to accommodate censoring effect
- Emura and Wang [2010]: weighted log rank statistic based on Copula

Existing tests

Tests for the Markov condition in illness-death model:



- de Uña-Álvarez [2012], Rodríguez-Girondo and Uña-Álvarez [2012]:
  - local tests based on marginal estimators and Kendall's tau
  - poor performance under moderate censoring.
- Rodríguez-Girondo and Uña-Álvarez [2016]:
  - double bootstrap for maximized local Kendall's tau
  - improved but still sensitive to censoring.

Permutation test

Permutation test

- In general, permutation test consists of the following procedures:
  1. Generate a large number of permuted data under null
  2. For each permuted data, compute a test statistics
  3. Compute a *p*-value
- The knowledge of the distribution of a test statistics under null hypothesis is not required

Permutation

- Permute T and let $\{(T_i^*, X_i); i = 1, \ldots, n\}$ be the permuted data.
- We consider two permutation approaches under left truncation
  1. Conditional permutation [Tsai, 1990, Efron and Petrosian, 1992]
  2. Unconditional permutation

Conditional permutation

- The conditional permutation procedure:
  1. Initialize with $m = 1$
  2. For X_m, selects a T_m^* from $\{i : T_i \leq X_i\}$
  3. Remove T_m^* from $\{T_1, \ldots, T_n\}$ and repeat step 2. with $m = 2, \ldots, n$.

Conditional permutation approach – example

- Suppose the observed data consists of 4 observations:

$$\{(X, T) : (3, 2), (5, 1), (8, 7), (9, 6)\}$$

- Four possible legal permutations:

$$\begin{array}{llll} \{(3,\ 1) & (5,\ 2) & (8,\ 6) & (9,\ 7)\} \\ \{(3,\ 1) & (5,\ 2) & (8,\ 7) & (9,\ 6)\} \\ \{(3,\ 2) & (5,\ 1) & (8,\ 6) & (9,\ 7)\} \\ \{(3,\ 2) & (5,\ 1) & (8,\ 7) & (9,\ 6)\} \end{array}$$

Unconditional permutation

- The unconditional permutation approach:
  1. Permutes T across all subjects in the dataset
  2. Delete those with $T_i^* > X_i, i = 1, \ldots, n$

Unconditional permutation approach – example

- Suppose the observed data consists of 4 observations:

$$\{(X, T) : (3, 2), (5, 1), (8, 7), (9, 6)\}$$

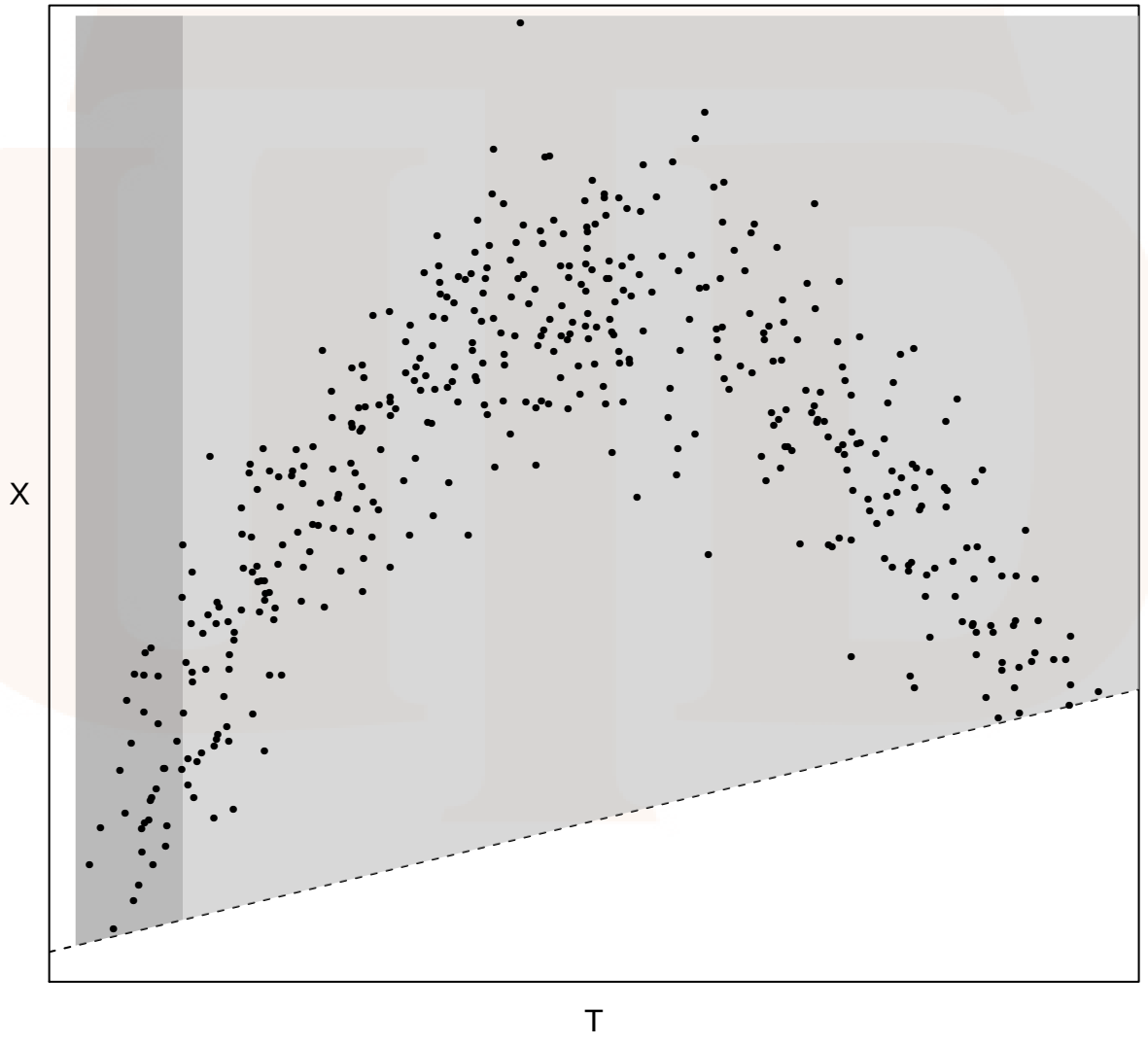
- Twenty four (4! = 24) possible legal permutations:

$$\begin{array}{llll}
\{(3,\ 1) & (5,\ 2) & (8,\ 6) & (9,\ 7)\} \\
\{(3,\ 1) & (5,\ 2) & (8,\ 7) & (9,\ 6)\} \\
\{(3,\ 1) & \cancel{(5,\ 6)} & (8,\ 2) & (9,\ 7)\} \\
\{(3,\ 1) & \cancel{(5,\ 6)} & (8,\ 7) & (9,\ 2)\}
\end{array}$$

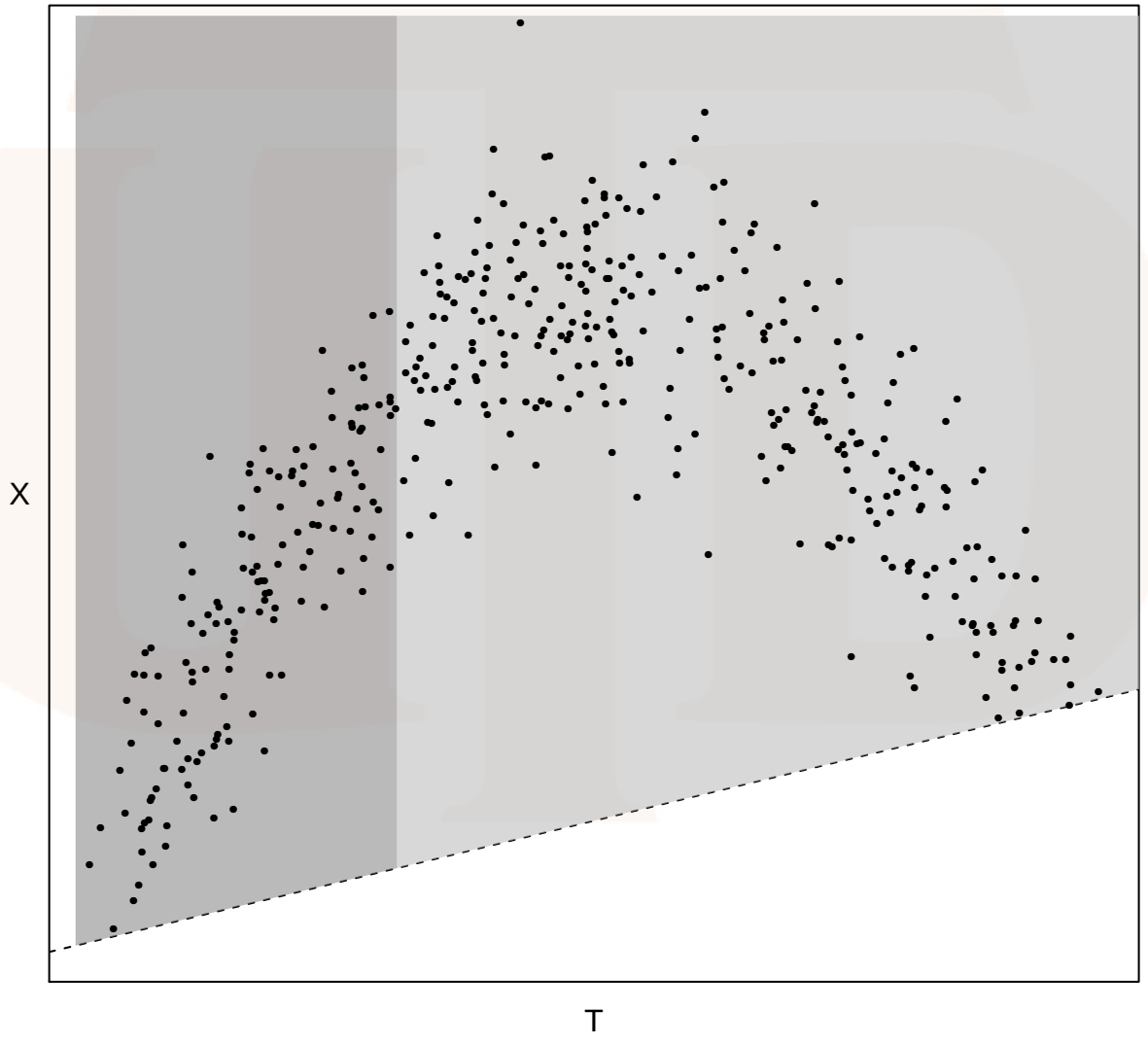
Test statistics: minimally selected p-value, minp_1

- Minimally selected p-value (minp) tests to detect non-monotone dependencies.
- We proposed to obtain minp_1 p-value from:
  1. Partition the data into two groups: $\{T < t\}$ or $\{T > t\}$
  2. Compute the log-rank p-value for the two groups
  3. Repeat 1. and 2. for $t \in \{T_1, \ldots, T_n\}$
  4. The minp_1 test statistic is the minimum of these p-values
- For every cut-point, require at least E events in the each group.

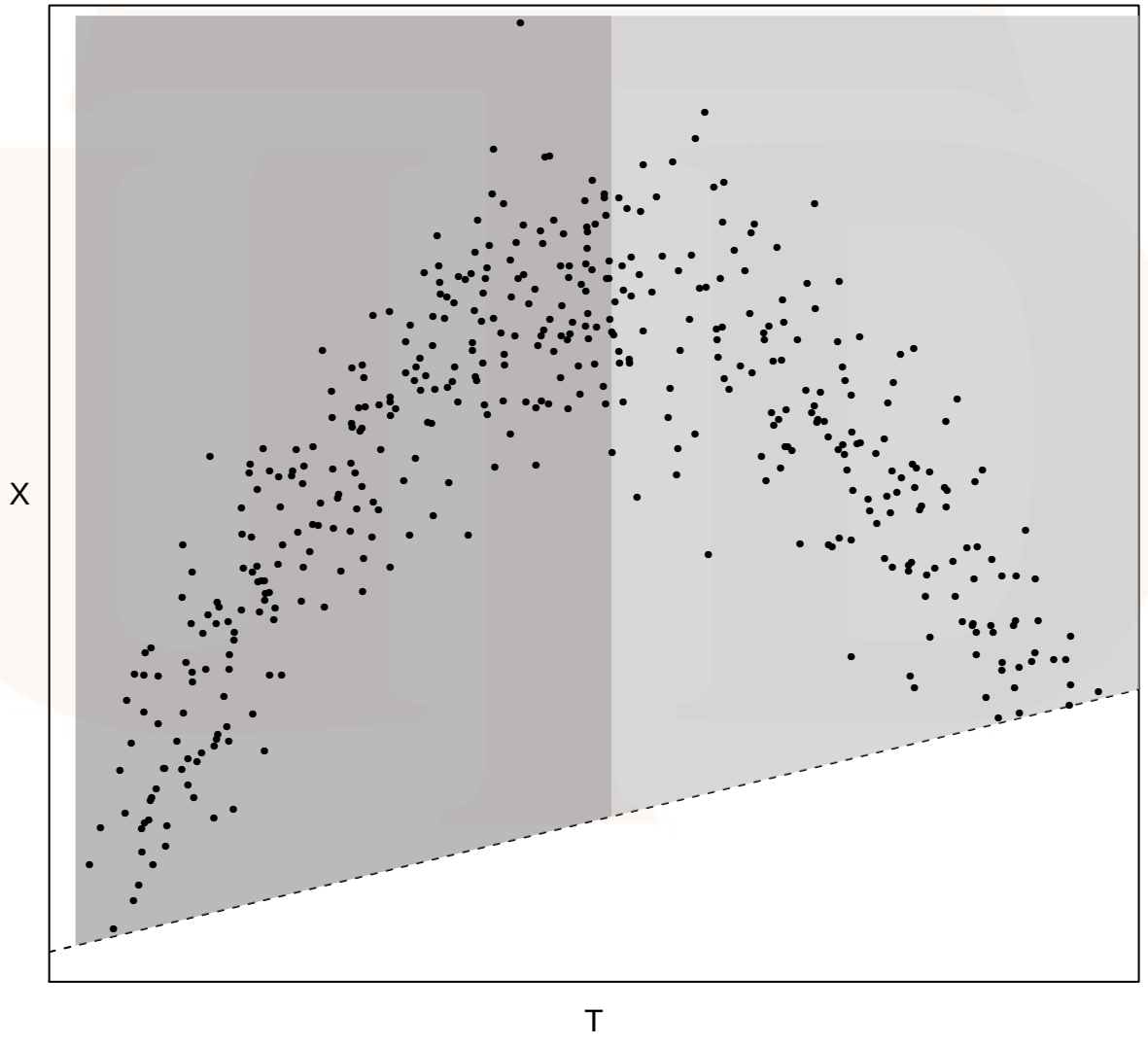
Minimally selected p-value, minp_1



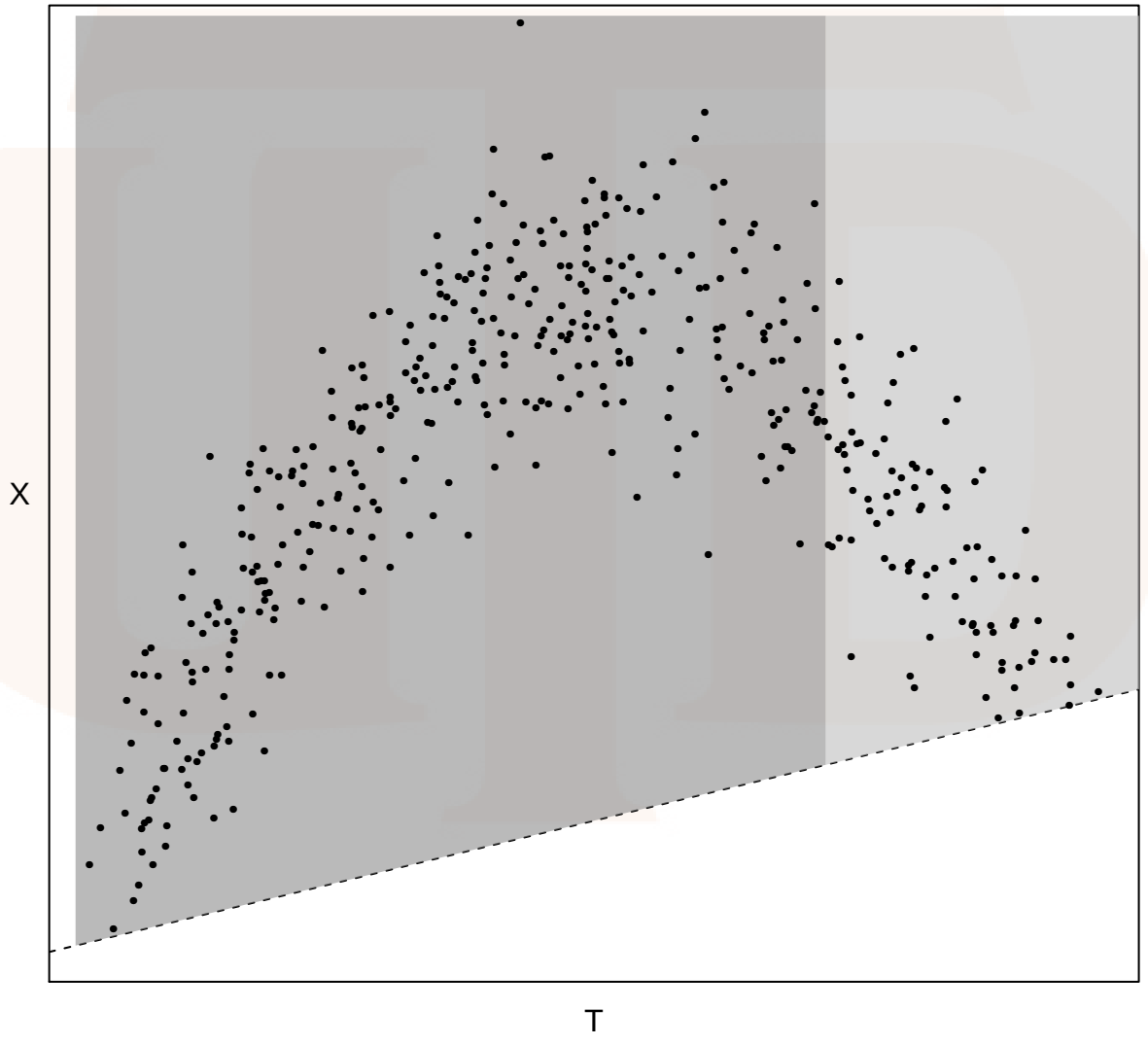
Minimally selected p-value, minp_1



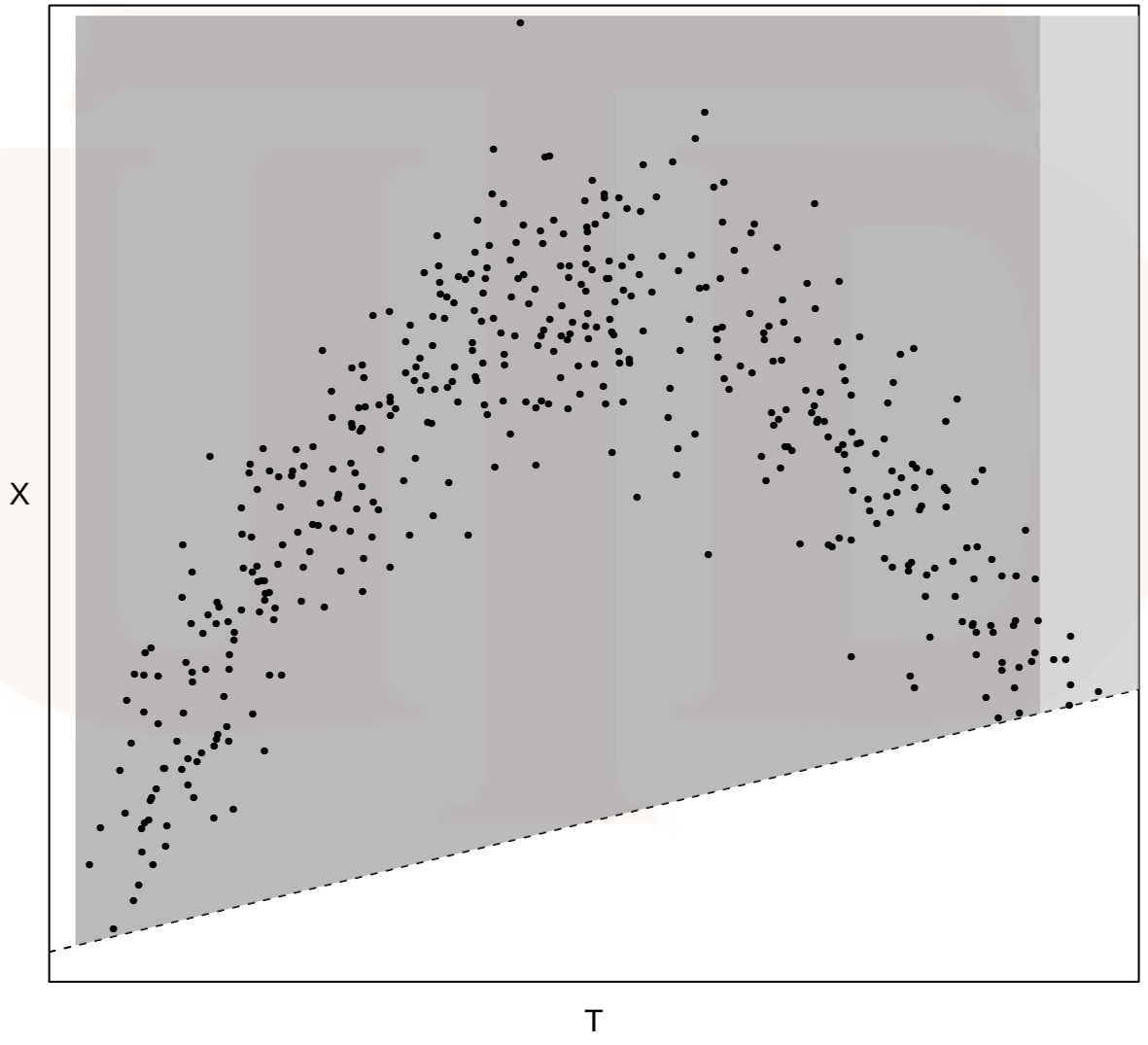
Minimally selected p-value, minp_1



Minimally selected p-value, minp_1



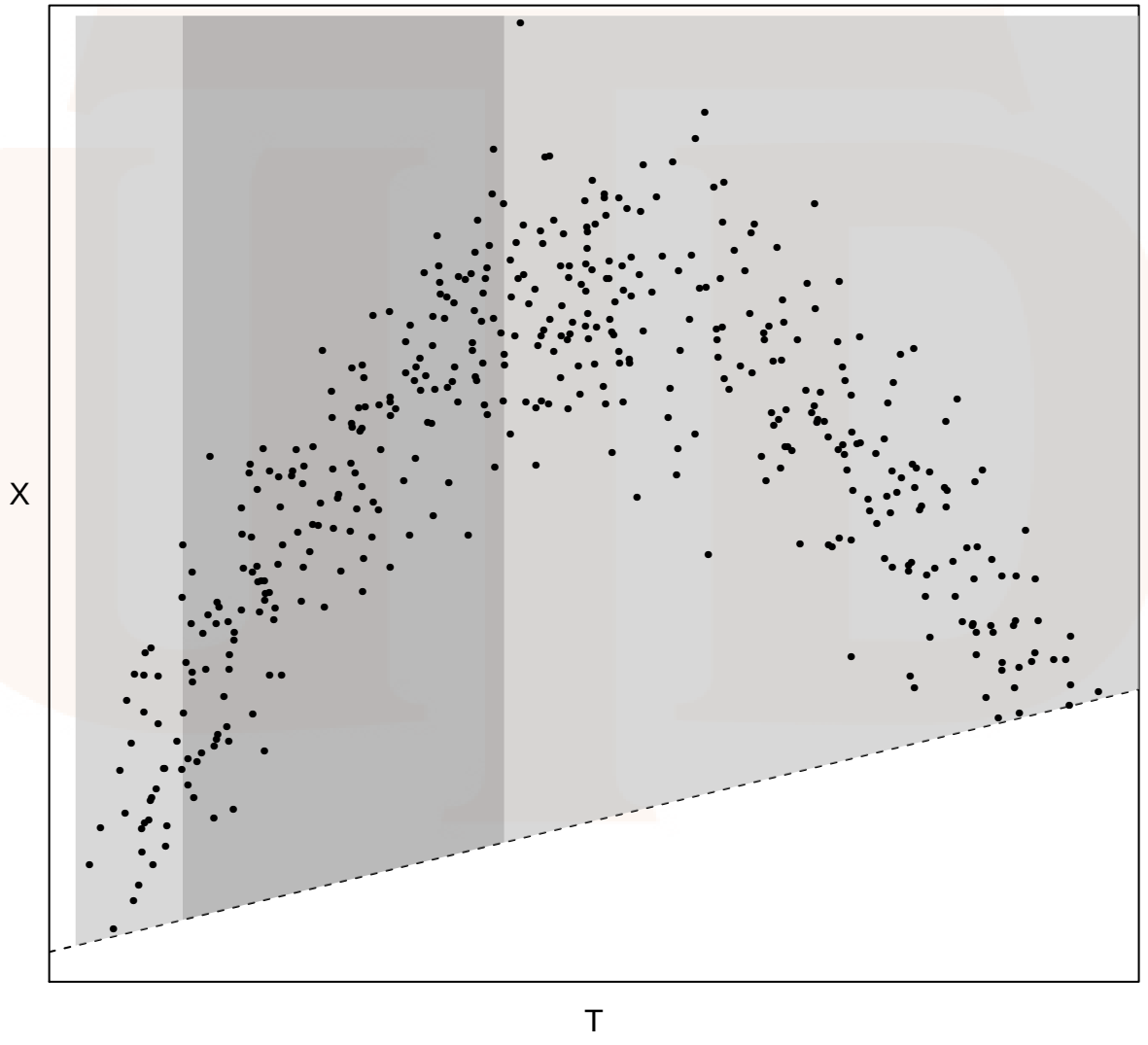
Minimally selected p-value, minp_1



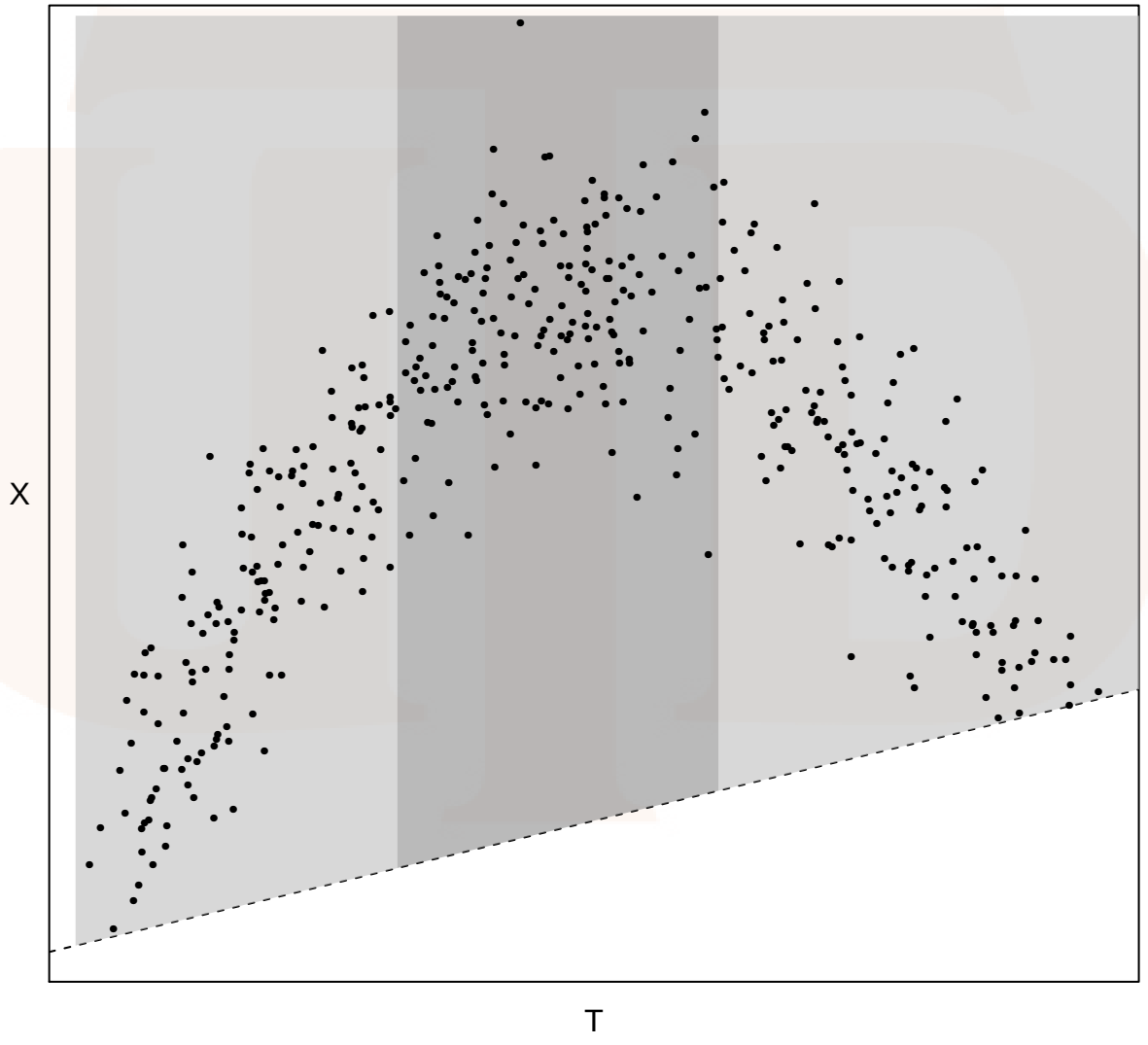
Minimally selected p-value, minp_2

- An alternative is the minp_2 test:
  1. Partition the data into two groups: $\{T \in (t-\epsilon, t+\epsilon)\}$ or $\{T \notin (t-\epsilon, t+\epsilon)\}$
  2. Compute the log-rank p-value for the two groups
  3. Repeat 1. and 2. for $t \in \{T_1, \ldots, T_n\}$
  4. The minp_2 test statistic is the minimum of these p-values
- This allows for X to be associated with moderate T differently from small or large T.
- Choose ϵ so that each group retains at least E events.

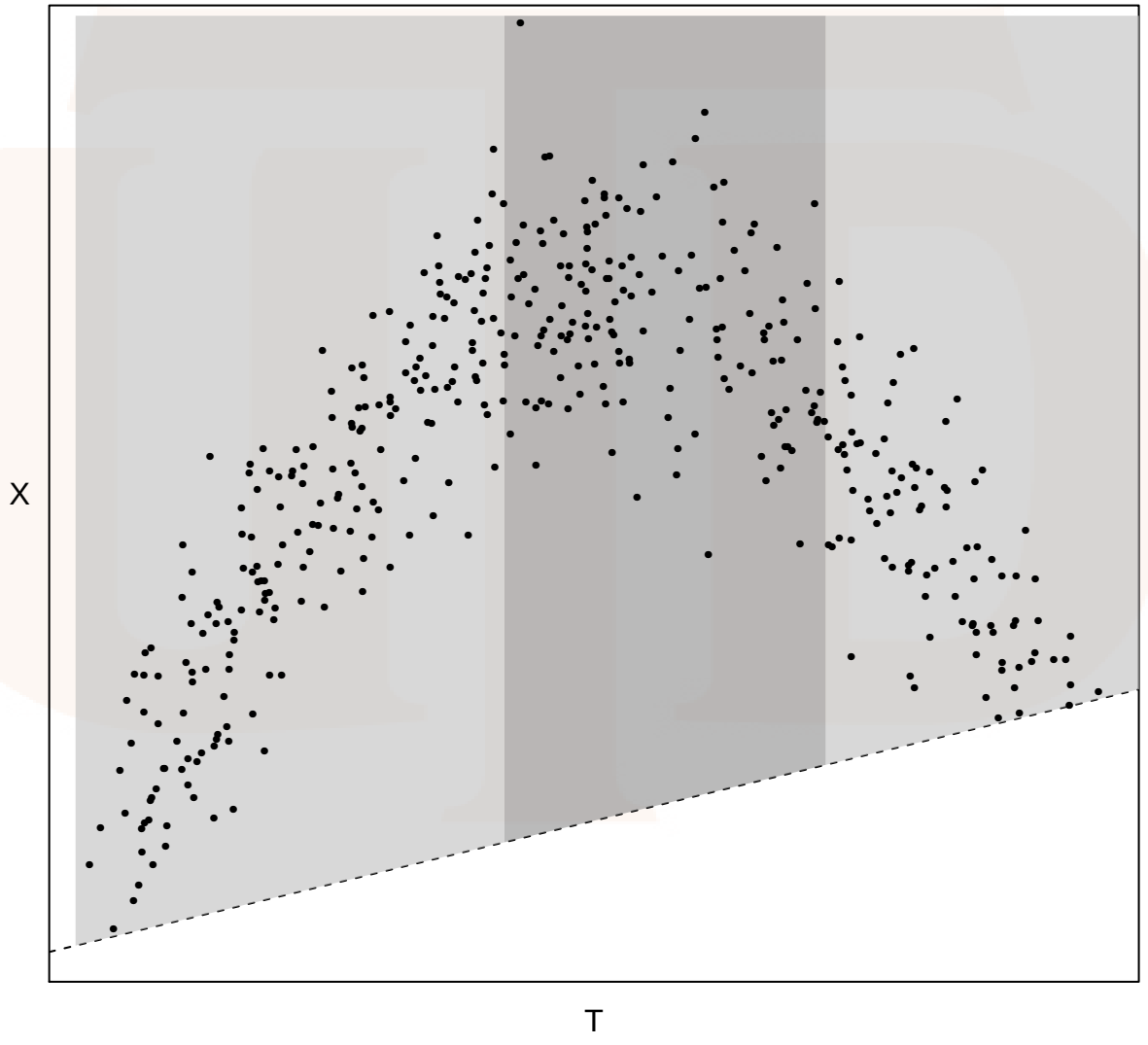
Minimally selected p-value, minp_2



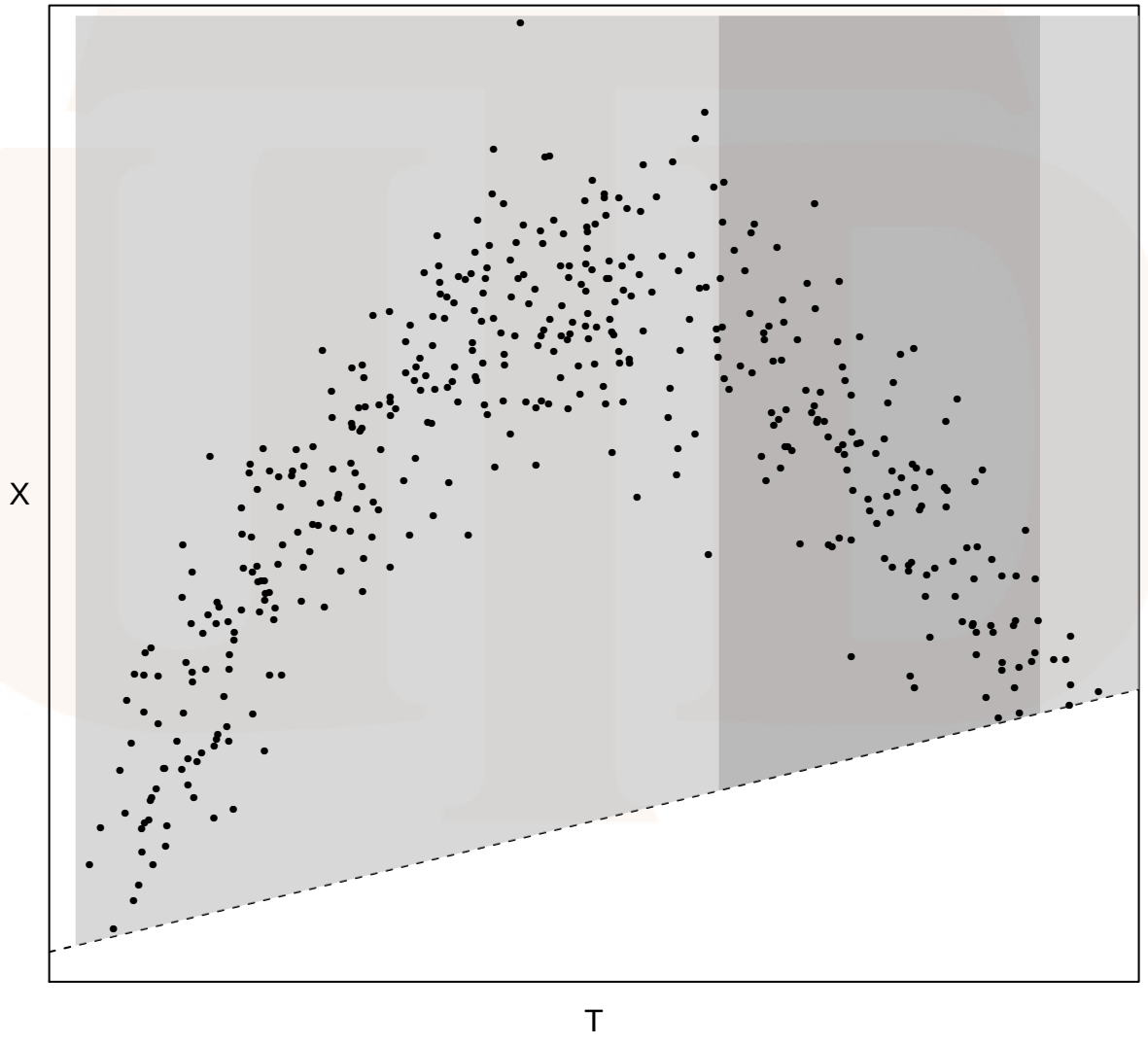
Minimally selected p-value, minp_2



Minimally selected p-value, minp_2



Minimally selected p-value, minp_2



Permutation p-values

- We approximate the p-value with

$$\frac{\sum_{i=1}^{N^*} I(|z_i^*| \geq |z_{\text{obs}}|) + 1}{N^* + 1},$$

where z_i^*'s and z_{obs} are the test statistics based on the sampled permutation and observed data, respectively.

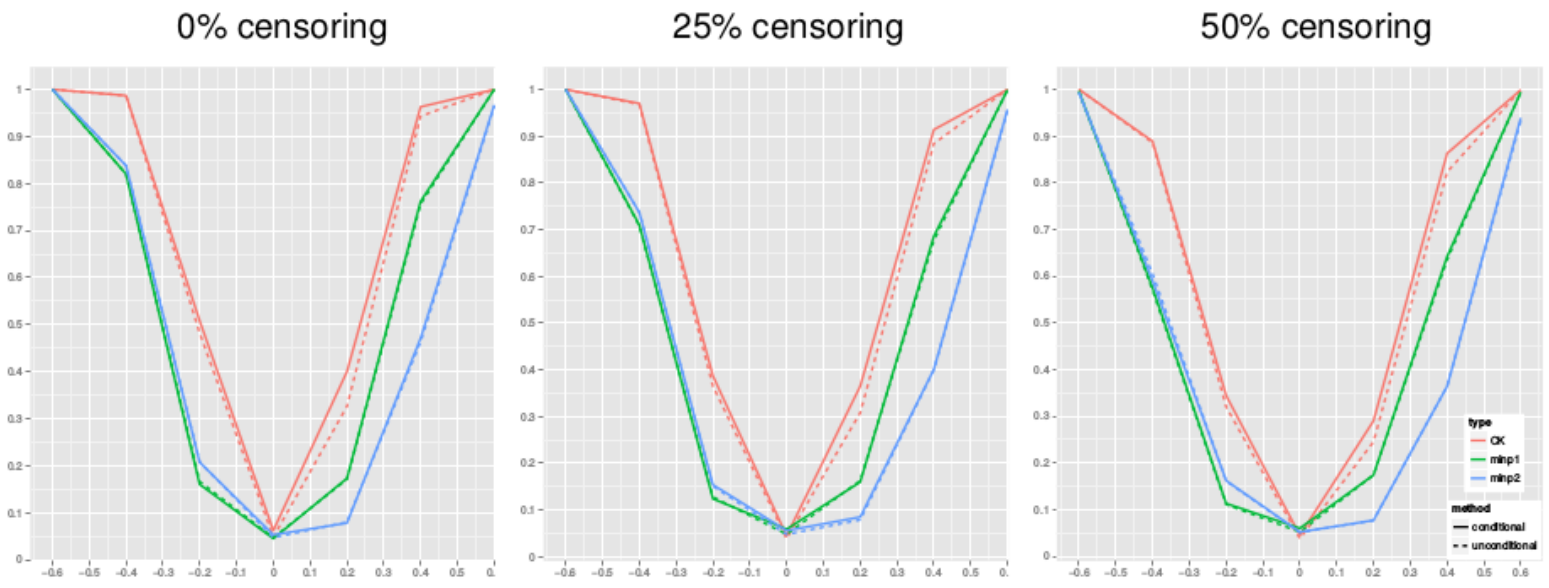
Simulation

Simulation 1: Monotone settings

- Generate (X, T) from a bivariate normal copula
  - $X \sim \text{Weibull}(3, 8.5)$
  - $T \sim \exp(0.2)$
- Nine levels of dependence
  - $\tau = 0, \pm 0.2, \pm 0.4, \pm 0.6, \pm 0.8$
- Sample size after truncation: 100 and 200.
- Censoring times follow an independent Uniform$(0, c)$
  - $0\%, 25\%$, and 50% after truncation
- 5000 permutations
- 1000 replications
- We compare the rejection proportions at a significant level of 0.05

Simulation 1: Rejection proportion

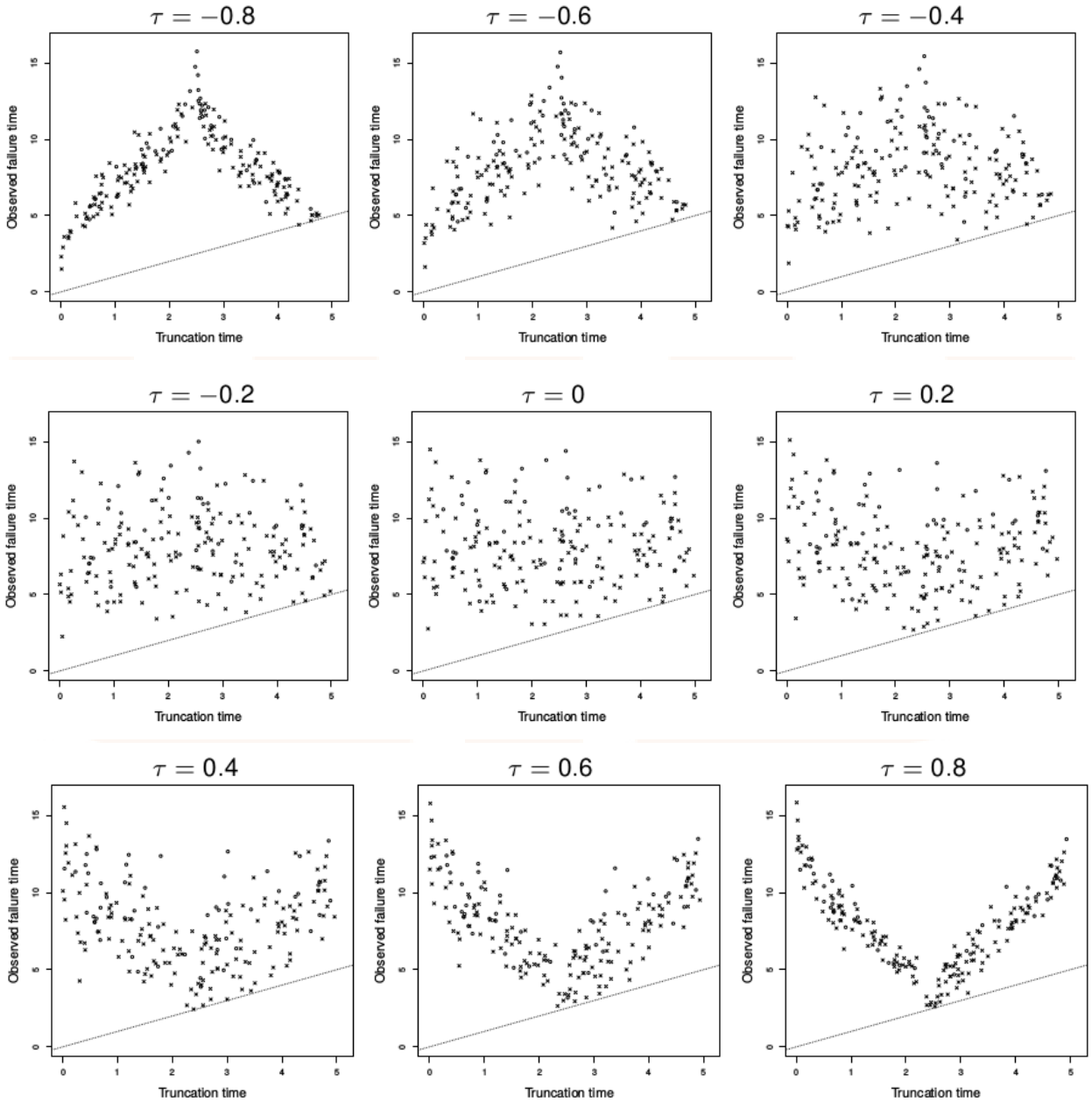
- Rejection proportion with $n = 100$:



Simulation 2: Non-monotone settings

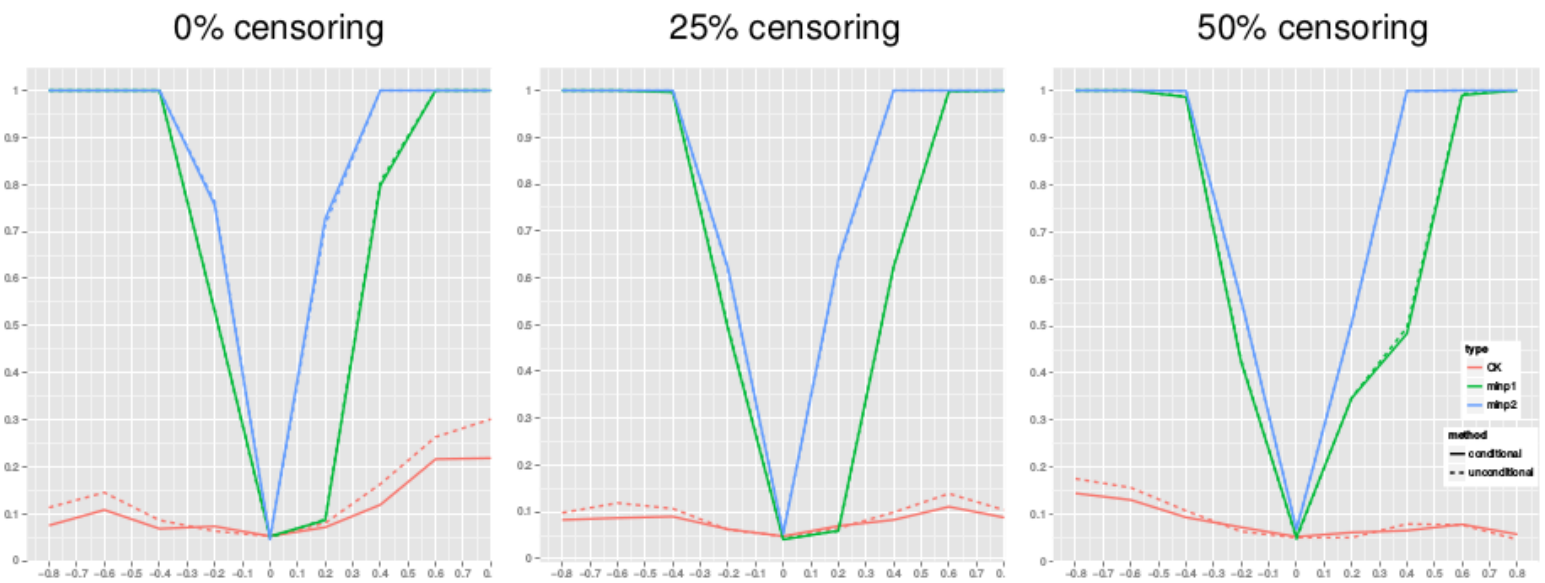
- Generate $(\mid T - 2.5 \mid, X)$ from a bivariate normal copula
  - $T \sim \text{Uniform}(0, 5)$
  - $X \sim \text{Weibull}(3, 8.5)$
- Nine dependence level
  - $0, \pm 0.2, \pm 0.4, \pm 0.6, \pm 0.8$
- Sample size after truncation: 100 and 200
- Censoring times follow an independent $\text{Uniform}(0, c)$
  - 0%, 25%, and 50% after truncation
- 5000 permutations
- 1000 replications
- We compare the rejection proportions at a significant level of 0.05.

Simulation 2: Data



Simulation 2: Rejection proportion

- Rejection proportion with $n = 100$:



Conclusion

Conclusion

- Minp tests are powerful against non-monotone dependencies.
- Minp tests are able to detect nonlinear dependence not detected by Kendall's tau in the aging study
- Unconditional permutation is a reasonable alternative
  - faster
  - useful for evaluation of statistics that fix risk set sizes
  - not strictly valid due to non-exchangeability
  - preserve type I error and have high power

References

References

M. D. Austin and R. A. Betensky. Eliminating bias due to censoring in Kendall's tau estimators for quasi-independence of truncation and failure. *Computational Statistics & Data Analysis*, 73: 16–26, 2014.

C.-H. Chen, W.-Y. Tsai, and W.-H. Chao. The product-moment correlation coefficient and linear regression for truncated data. *Journal of the American Statistical Association*, 91(435): 1181–1186, 1996.

J. de Uña-Álvarez. On the Markov three-state progressive model. In *Recent Advances in System Reliability*, pages 269–281. Springer, 2012.

B. Efron and V. Petrosian. A simple test of independence for truncated data with applications to redshift surveys. *The Astrophysical Journal*, 399:345–352, 1992.

T. Emura and W. Wang. Testing quasi-independence for truncation data. *Journal of Multivariate Analysis*, 101(1):223–239, 2010.

M. P. Jones and J. Crowley. Nonparametric tests of the Markov model for survival data. *Biometrika*, 79(3):513–522, 1992.

E. C. Martin and R. A. Betensky. Testing quasi-independence of failure and truncation times via conditional Kendall's tau. *Journal of the American Statistical Association*, 100(470):484–492, 2005. doi: 10.1198/016214504000001538.

M. Rodríguez-Girondo and J. Uña-Álvarez. A nonparametric test for Markovianity in the illness-death model. *Statistics in Medicine*, 31(30):4416–4427, 2012.

M. Rodríguez-Girondo and J. d. Uña-Álvarez. Methods for testing the Markov condition in the illness–death model: A comparative study. *Statistics in Medicine*, 35(20):3549–3562, 2016.

W.-Y. Tsai. Testing the assumption of independence of truncation time and failure time. *Biometrika*, 77(1):169–177, 1990.

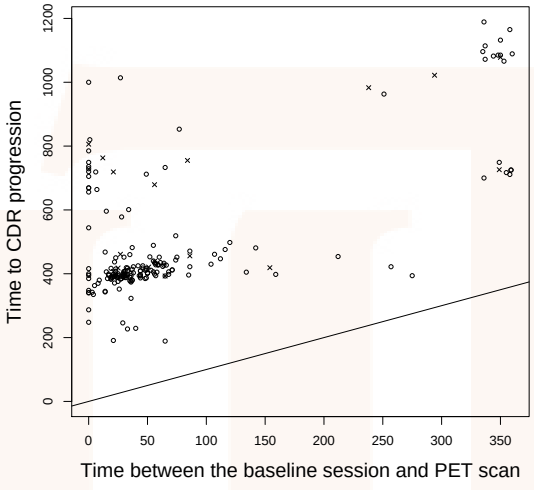
B. Vakulenko-Lagun, J. Qian, S. Chiou, and R. Betensky. Nonidentifiability in the presence of factorization for truncated data. *Biometrika*, 2018+.

Aging study

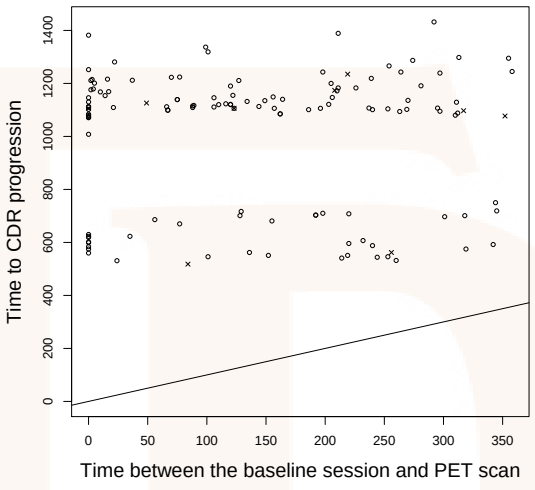
- Cognitively normal older individuals (n = 490):
  1. Alzheimer's Disease Neuroimaging Initiative (n = 198)
  2. Australian Imaging Biomarkers and Lifestyle Study of Ageing (n = 131)
  3. Harvard Aging Brain Study (n = 161)
- Participants had a Clinical Dementia Rating (CDR) 0 at enrollment
- Test for quasi-independence between time to PET scan and time to CDR progression

Aging study

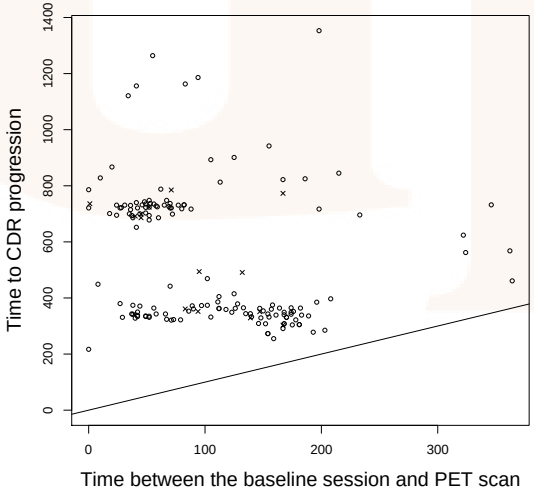
ADNI



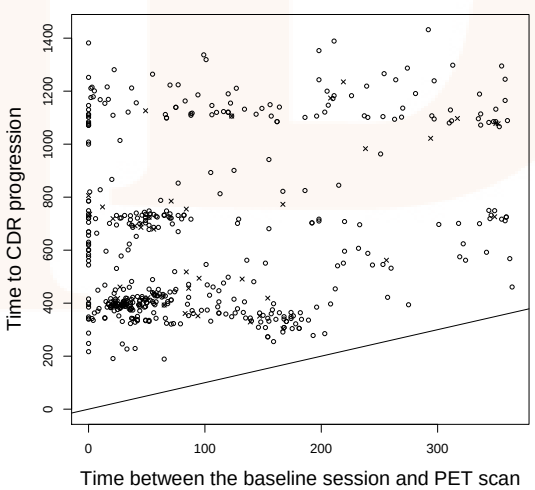
AIBL



HABS



ALL



Aging study – results ($E = 8$, 5000 permutations)

					Conditional		Unconditional	
	n	$\Delta = 1$	$\hat{\tau}_c^*$	ACK	minp_1	minp_2	minp_1	minp_2
ALL	490	41	−0.010	0.904	0.518	0.410	0.507	0.394
ADNI	198	17	0.149	0.292	0.111	0.019	0.082	0.018
HABS	161	15	−0.174	0.276	0.703	0.076	0.692	0.074
AIBL	131	9	−0.185	0.418	0.456	0.317	0.257	0.308

- no evidence of monotone dependence
- minp_2 does appear to be more powerful than minp_1
- Some evidence of nonlinear association in ADNI; possibly in HABS (marginally significant minp_2)