Benefit-risk assessment via Generalized Pairwise Comparisons

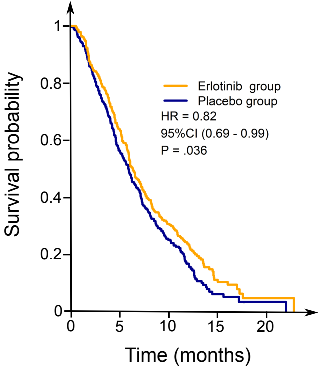
Brice Ozenne[1,2], Esben Budtz-Jørgensen[1], Julien Péron[3,4]

1. Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark
2. Neurobiology Research Unit, University Hospital of Copenhagen, Copenhagen, Denmark
3. Institut de Cancérologie des Hospices Civils de Lyon, Lyon, France
4. Laboratoire de Biométrie et Biologie Evolutive, Equipe Biostatistique-Santé, CNRS UMR 5558, Université Claude Bernard Lyon 1, Villeurbanne, France.

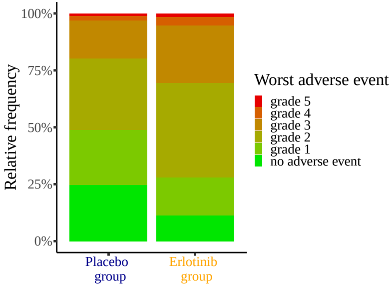
December 18th, 2022 - CMstatistics

Clinical trials in oncology - Moore et al., 2007

Benefit
(time to death)



Risk
(non-lethal adverse event)



Benefit risk assessment

I do not think there is a good objective approach.

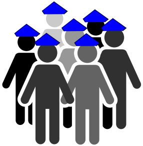
- outcome-specific analyses are not sufficient

What about a good subjective approach?

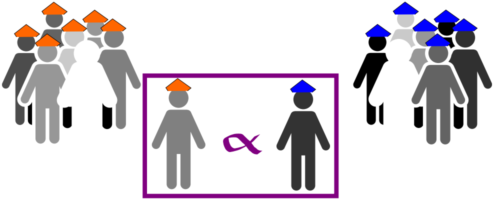
Patient OB preference

1. gain in survival of at least 2 months
2. otherwise, least serious adverse event

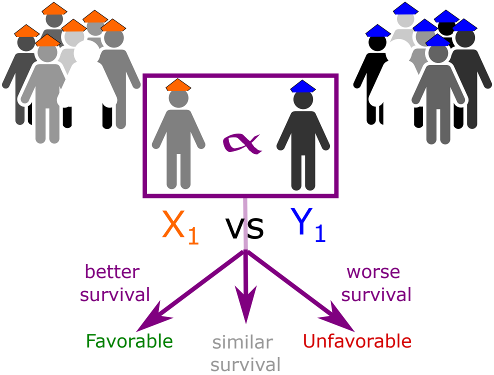
Generalized Pairwise Comparisons (GPC)



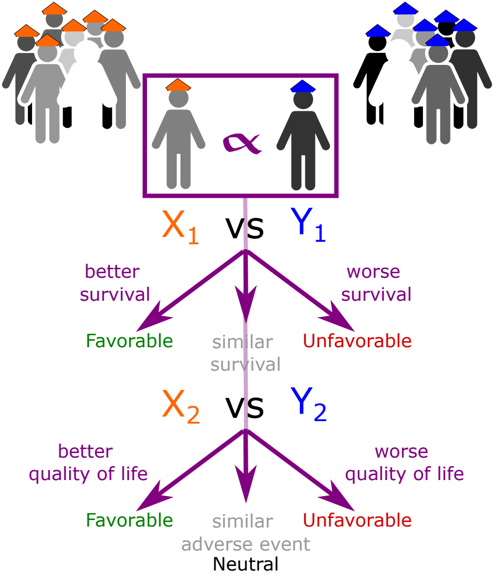
Generalized Pairwise Comparisons (GPC)



Generalized Pairwise Comparisons (GPC)



Generalized Pairwise Comparisons (GPC)



Parameter of interest

Net benefit:

$$\Delta = \mathbb{P}[X_1 \geq Y_1 + \tau_1] - \mathbb{P}[Y_1 \geq X_1 + \tau_1]$$
$$+ \mathbb{P}[X_2 \geq Y_2 + \tau_2, |X_1 - Y_1| < \tau_1] - \mathbb{P}[Y_2 \geq X_2 + \tau_2, |X_1 - Y_1| < \tau_1]$$
$$= U_1^+ - U_1^- + U_2^+ - U_2^-$$

where:

- $\boldsymbol{\tau} = (\tau_1 = 2, \tau_2 = 0.1)$: smallest clinically relevant difference
- $\boldsymbol{X} = (X_1, X_2)$: outcomes in the experimental arm
- $\boldsymbol{Y} = (Y_1, Y_2)$: outcomes in the placebo arm

Parameter of interest

Net benefit:

$$\begin{aligned}\Delta &= \mathbb{P}[X_1 \geq Y_1 + \tau_1] - \mathbb{P}[Y_1 \geq X_1 + \tau_1] \\ &+ \mathbb{P}[X_2 \geq Y_2 + \tau_2, |X_1 - Y_1| < \tau_1] - \mathbb{P}[Y_2 \geq X_2 + \tau_2, |X_1 - Y_1| < \tau_1] \\ &= U_1^+ - U_1^- + U_2^+ - U_2^-\end{aligned}$$

where:

- $\boldsymbol{\tau} = (\tau_1 = 2, \tau_2 = 0.1)$: smallest clinically relevant difference
- $\boldsymbol{X} = (X_1, X_2)$: outcomes in the experimental arm
- $\boldsymbol{Y} = (Y_1, Y_2)$: outcomes in the placebo arm

$$\Delta = \begin{cases} 1, \text{ treatment always better} \\ 0, \text{ no difference} \\ -1, \text{ treatment always worse} \end{cases}$$

Parameter of interest

Net benefit:

$$\Delta = \mathbb{P}[X_1 \geq Y_1 + \tau_1] - \mathbb{P}[Y_1 \geq X_1 + \tau_1]$$
$$+ \mathbb{P}[X_2 \geq Y_2 + \tau_2, |X_1 - Y_1| < \tau_1] - \mathbb{P}[Y_2 \geq X_2 + \tau_2, |X_1 - Y_1| < \tau_1]$$
$$= U_1^+ - U_1^- + U_2^+ - U_2^-$$

where:

- $\boldsymbol{\tau} = (\tau_1 = 2, \tau_2 = 0.1)$: smallest clinically relevant difference
- $\boldsymbol{X} = (X_1, X_2)$: outcomes in the experimental arm
- $\boldsymbol{Y} = (Y_1, Y_2)$: outcomes in the placebo arm

$$\Delta = \begin{cases} 1, \text{ treatment always better} \\ 0, \text{ no difference} \\ -1, \text{ treatment always worse} \end{cases}$$

Estimation

- U-statistic
  - handling right-censoring

Notations & assumptions

Consider the following two samples:

- $(\boldsymbol{x}_i)_{i=1}^m = (\tilde{x}_{1i}, \varepsilon_{1i}, x_{2i})_{i=1}^m$
- $\left(\boldsymbol{y}_j\right)_{j=1}^n = (\tilde{y}_{1j}, \eta_{1j}, y_{2j})_{j=1}^n$

i.e. (survival time, censoring indicator, categorical outcome)

[1] independent and identically distributed

Notations & assumptions

Consider the following two samples:

- $(\boldsymbol{x}_i)_{i=1}^{m} = (\tilde{x}_{1i}, \varepsilon_{1i}, x_{2i})_{i=1}^{m}$
- $\left(\boldsymbol{y}_j\right)_{j=1}^{n} = (\tilde{y}_{1j}, \eta_{1j}, y_{2j})_{j=1}^{n}$

i.e. (survival time, censoring indicator, categorical outcome)

Assumptions:

- independent samples
- observations iid[1] within sample
- ratio $\frac{m}{n} \to \rho \in]0, 1[$ when $m + n \to \infty$.
- right-censoring independent of the outcome within sample

[1] independent and identically distributed

Estimation

With complete data we can estimate:

$$U_1^+ = \mathbb{P}\left[X_1 \geq Y_1 + \tau_1\right]$$

using a two-sample U-statistic:

$$\widehat{U}_1^+ = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1}$$

where $\mathbb{1}_\bullet$ is the indicator function: 1 if $\bullet$ is true, 0 otherwise.

Estimation

With complete data we can estimate:

$$U_1^+ = \mathbb{P}[X_1 \geq Y_1 + \tau_1]$$

using a two-sample U-statistic:

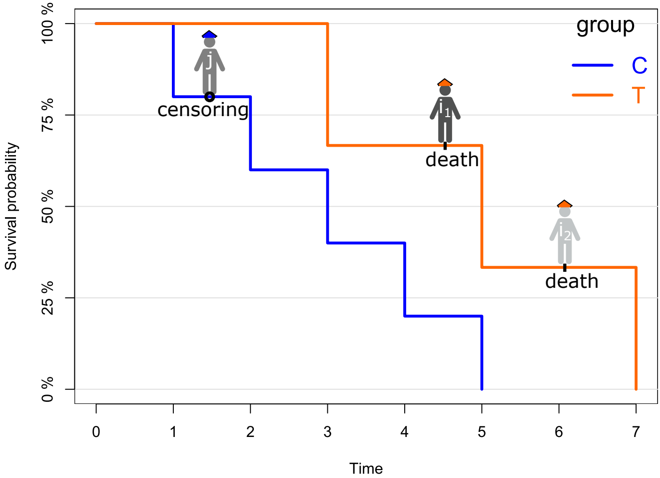
$$\widehat{U}_1^+ = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1}$$

where $\mathbb{1}_\bullet$ is the indicator function: 1 if $\bullet$ is true, 0 otherwise.

In presence of right censoring:

- inverse probability of censoring weights e.g. Zhang et al. (2022)
- "Peron scoring rule" Péron et al. (2018)

Peron scoring rule (intuition)



$$\mathbb{E}\left[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} \middle| \tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}\right] = 0.75 \text{ for } i_1$$

$$= 1 \text{ for } i_2$$

Estimation with right censoring

Introducing S_X and S_Y the group-specific survival curves:

$$\begin{aligned}\widehat{U}_1^+ =& \widehat{U}_1^+(S_X, S_Y)\\ =& \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbb{E}\left[\mathbb{1}_{x_{1i}\geq y_{1j}+\tau_1}\middle|\tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, S_X, S_Y\right]\end{aligned}$$

Estimation with right censoring

Introducing S_X and S_Y the group-specific survival curves:

$$\begin{aligned}\widehat{U}_1^+ &= \widehat{U}_1^+(S_X, S_Y) \\ &= \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbb{E}\left[\mathbb{1}_{x_{1i}\geq y_{1j}+\tau_1}|\tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, S_X, S_Y\right]\end{aligned}$$

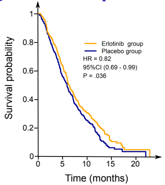
Example: $\varepsilon_{1i} = 1$ (event), $\eta_{1j} = 0$ (censored), $x_{1i} \geq \tilde{y}_{1j} + \tau$

$$\mathbb{E}\left[\mathbb{1}_{x_{1i}\geq y_{1j}+\tau_1}|\tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, S_X, S_Y\right] = 1 - \frac{S_Y(x_{1i} - \tau_1)}{S_Y(\tilde{y}_{1j})}$$

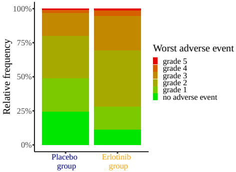
$$= 1 - \frac{0.2}{0.8} = 0.75 \text{ (for } i_1 \text{ vs. } j)$$

Back to the motivating example

- about 15% of right-censoring



- S_X and S_Y are estimated using Kaplan-Meier (denoted $\widehat{S}_X$ and $\widehat{S}_Y$)



Priority	Favorable	Unfavorable	Neutral	$\widehat{\delta}$
1 (survival, 2 months)	42.0 %	33.5 %	24.5%	8.5%
2 (adverse event)	6.8 %	11.9 %	5.9%	-5.1%

Overall: $\widehat{\Delta} = 3.4\%$

Uncertainty quantification

- with complete data
- with right-censoring

Intuition

$$\widehat{U}_1^+ = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbb{1}_{x_{1i}\geq y_{1j}+\tau_1} \text{ is an average !}$$

Intuition

$$\widehat{U}_1^+ = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} \text{ is an average !}$$

⚠ $\left(\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1}\right)_{i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}}$ are not independent
e.g. $\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1}$ and $\mathbb{1}_{x_{1i} \geq y_{1j'} + \tau_1}$ both depends on x_{1i}

Intuition

$$\widehat{U}_1^+ = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbb{1}_{x_{1i}\geq y_{1j}+\tau_1} \text{ is an average !}$$

⚠ $\left(\mathbb{1}_{x_{1i}\geq y_{1j}+\tau_1}\right)_{i\in\{1,\ldots,n\},j\in\{1,\ldots,m\}}$ are not independent
e.g. $\mathbb{1}_{x_{1i}\geq y_{1j}+\tau_1}$ and $\mathbb{1}_{x_{1i}\geq y_{1j'}+\tau_1}$ both depends on x_{1i}

This motivate the following H-decomposition (Lee, 1990):

$$\widehat{U}_1^+ - U_1^+ = \frac{1}{m}\sum_{i=1}^{m}H_i^{(1,0)} + \frac{1}{n}\sum_{j=1}^{n}H_j^{(0,1)} + \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}H_{ij}^{(1,1)}$$

- sum of uncorrelated U-statistics of increasing order
- with variance of decreasing order in n, m.

Asymptotic distribution of $\widehat{U}_1^+$

$$\widehat{U}_1^+ - U_1^+ = \frac{1}{m}\sum_{i=1}^{m} H_i^{(1,0)} + \frac{1}{n}\sum_{j=1}^{n} H_j^{(0,1)} + \underbrace{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} H_{ij}^{(1,1)}}_{\text{asymptotically neglectable}}$$

$$H_i^{(1,0)} = \mathbb{E}[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | x_{1i}] - U_1^+$$

$$H_j^{(0,1)} = \mathbb{E}[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | y_{1j}] - U_1^+$$

Asymptotic distribution of $\widehat{U}_1^+$

$$\widehat{U}_1^+ - U_1^+ = \frac{1}{m}\sum_{i=1}^{m} H_i^{(1,0)} + \frac{1}{n}\sum_{j=1}^{n} H_j^{(0,1)} + \underbrace{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} H_{ij}^{(1,1)}}_{\text{asymptotically neglectable}}$$

$$H_i^{(1,0)} = \mathbb{E}[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | x_{1i}] - U_1^+$$
$$H_j^{(0,1)} = \mathbb{E}[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | y_{1j}] - U_1^+$$

- we have means of independent terms! So by the central limit theorm (CLT), $\widehat{U}_1^+$ is normally distributed

$$\mathbb{V}ar\left[\widehat{U}_1^+\right] \approx \frac{1}{m^2}\sum_{i=1}^{m}\left(H_i^{(1,0)}\right)^2 + \frac{1}{n^2}\sum_{j=1}^{n}\left(H_j^{(0,1)}\right)^2$$

Asymptotic distribution of $\widehat{U}_1^+$

$$\widehat{U}_1^+ - U_1^+ = \frac{1}{m}\sum_{i=1}^{m} H_i^{(1,0)} + \frac{1}{n}\sum_{j=1}^{n} H_j^{(0,1)} + \underbrace{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} H_{ij}^{(1,1)}}_{\text{asymptotically neglectable}}$$

$$H_i^{(1,0)} = \mathbb{E}[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | x_{1i}] - U_1^+$$
$$H_j^{(0,1)} = \mathbb{E}[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | y_{1j}] - U_1^+$$

- we have means of independent terms! So by the central limit theorm (CLT), $\widehat{U}_1^+$ is normally distributed

$$\mathbb{V}ar\left[\widehat{U}_1^+\right] \approx \frac{1}{m^2}\sum_{i=1}^{m}\left(H_i^{(1,0)}\right)^2 + \frac{1}{n^2}\sum_{j=1}^{n}\left(H_j^{(0,1)}\right)^2$$

- similar arguments hold for $\widehat{\Delta}$

Simulation results (1/2)

- single time to event outcome, 20 000 repetitions
- confidence intervals computed after `tanh` transform and backtransformed (ensure bounds within [-1;1])

Here $\tau = 0.5$ is a large threshold (approx. median survival time)

What about right-censoring

If we knew the survival curves:

- Re-use the H-decomposition with $\mathbb{E}\left[\mathbb{1}_{x_{1i} \geq y_{1j}+\tau_1} \middle| \tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, S_X, S_Y\right]$ instead of $\mathbb{1}_{x_{1i} \geq y_{1j}+\tau_1}$

What about right-censoring

If we knew the survival curves:

- Re-use the H-decomposition with $\mathbb{E}\left[\mathbb{1}_{x_{1i} \geq y_{1j}+\tau_1} \middle| \tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, S_X, S_Y\right]$ instead of $\mathbb{1}_{x_{1i} \geq y_{1j}+\tau_1}$

⚠ Solely doing when estimating the survival leads to:

- correlated terms in the H-decomposition ($H_i^{(1,0)}$ and $H_{i\prime}^{(1,0)}$ may both depend on $\widehat{S}_X$ or $\widehat{S}_Y$).
- underestimation of the uncertainty

What about right-censoring

If we knew the survival curves:

- Re-use the H-decomposition with $\mathbb{E}\left[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} \middle| \tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, S_X, S_Y\right]$ instead of $\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1}$

⚠ Solely doing when estimating the survival leads to:

- correlated terms in the H-decomposition ($H_i^{(1,0)}$ and $H_{i'}^{(1,0)}$ may both depend on $\widehat{S}_X$ or $\widehat{S}_Y$).
- underestimation of the uncertainty

A decomposition in independent terms uses (Randles, 1982):

$$\widehat{U}_1^+(\widehat{S}_X, \widehat{S}_Y) - U_1^+(S_X, S_Y) = \underbrace{\widehat{U}_1^+(\widehat{S}_X, \widehat{S}_Y) - \widehat{U}_1^+(S_X, S_Y)}_{\text{new decomposition}} + \underbrace{\widehat{U}_1^+(S_X, S_Y) - U_1^+(S_X, S_Y)}_{\text{previous H-projection}}$$

Simplified survival model

Exponential model:

- $\widehat{S}_X(t) = \exp(-\widehat{\lambda}_X t)$ and $\widehat{S}_Y(t) = \exp(-\widehat{\lambda}_Y t)$
  where $\widehat{\lambda}_X = \frac{\sum_{i=1}^{m} \varepsilon_{1i}}{\sum_{i=1}^{m} \tilde{x}_{1i}} = \frac{\#\text{ death}}{\#\text{ follow-up time}}$ and $\widehat{\lambda}_Y = \frac{\sum_{j=1}^{n} \eta_{1j}}{\sum_{j=1}^{n} \tilde{y}_{1j}}$

Simplified survival model

Exponential model:

- $\widehat{S}_X(t) = \exp(-\widehat{\lambda}_X t)$ and $\widehat{S}_Y(t) = \exp(-\widehat{\lambda}_Y t)$ where $\widehat{\lambda}_X = \frac{\sum_{i=1}^m \varepsilon_{1i}}{\sum_{i=1}^m \tilde{x}_{1i}} = \frac{\text{\# death}}{\text{\# follow-up time}}$ and $\widehat{\lambda}_Y = \frac{\sum_{j=1}^n \eta_{1j}}{\sum_{j=1}^n \tilde{y}_{1j}}$

Example: $\varepsilon_{1i} = 1$, $\eta_{1j} = 0$, $x_{1i} \geq \tilde{y}_{1j} + \tau$:

$$\mathbb{E}\left[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} \middle| \tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, \widehat{S}_X, \widehat{S}_Y\right] = 1 - \frac{\widehat{S}_Y(x_{1i} - \tau_1)}{\widehat{S}_Y(\tilde{y}_{1j})}$$

$$= 1 - \exp\left(-\widehat{\lambda}_Y(x_{1i} - \tau_1 - \tilde{y}_{1j})\right)$$

Simplified survival model

Exponential model:

- $\widehat{S}_X(t) = \exp(-\widehat{\lambda}_X t)$ and $\widehat{S}_Y(t) = \exp(-\widehat{\lambda}_Y t)$
  where $\widehat{\lambda}_X = \frac{\sum_{i=1}^m \varepsilon_{1i}}{\sum_{i=1}^m \tilde{x}_{1i}} = \frac{\#\text{ death}}{\#\text{ follow-up time}}$ and $\widehat{\lambda}_Y = \frac{\sum_{j=1}^n \eta_{1j}}{\sum_{j=1}^n \tilde{y}_{1j}}$

Example: $\varepsilon_{1i} = 1$, $\eta_{1j} = 0$, $x_{1i} \geq \tilde{y}_{1j} + \tau$:

$$\mathbb{E}\left[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | \tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, \widehat{S}_X, \widehat{S}_Y\right] = 1 - \frac{\widehat{S}_Y(x_{1i} - \tau_1)}{\widehat{S}_Y(\tilde{y}_{1j})}$$

$$= 1 - \exp\left(-\widehat{\lambda}_Y(x_{1i} - \tau_1 - \tilde{y}_{1j})\right)$$

Estimates admit an iid decomposition, e.g.:

$$\widehat{\lambda}_Y - \lambda_Y = \frac{1}{n} \sum_{j=1}^n \frac{\lambda_Y}{\frac{1}{n} \sum_{j=1}^n \eta_{1j}} (\eta_{1j} - \tilde{y}_{1j} \lambda_Y) + o_p\left(\frac{1}{\sqrt{n}}\right)$$

New decomposition

We can use a first order Taylor expansion[2]:

$$\widehat{U}_1^+(\widehat{S}_X, \widehat{S}_Y) - \widehat{U}_1^+(S_X, S_Y) = \widehat{U}_1^+(\widehat{\lambda}_X, \widehat{\lambda}_Y) - \widehat{U}_1^+(\lambda_X, \lambda_Y)$$

$$= \frac{\partial \widehat{U}_1^+(\lambda_X, \lambda_Y)}{\partial \lambda_X}(\widehat{\lambda}_X - \lambda_X) + \frac{\partial \widehat{U}_1^+(\lambda_X, \lambda_Y)}{\partial \lambda_Y}(\widehat{\lambda}_Y - \lambda_Y) + o_p\left(\frac{1}{\sqrt{m}}\right) + o_p\left(\frac{1}{\sqrt{n}}\right)$$

[2] under smoothness conditions satisfied for the exponential model

New decomposition

We can use a first order Taylor expansion[2]:

$$\widehat{U}_1^+(\widehat{S}_X, \widehat{S}_Y) - \widehat{U}_1^+(S_X, S_Y) = \widehat{U}_1^+(\widehat{\lambda}_X, \widehat{\lambda}_Y) - \widehat{U}_1^+(\lambda_X, \lambda_Y)$$

$$= \frac{\partial \widehat{U}_1^+(\lambda_X, \lambda_Y)}{\partial \lambda_X}(\widehat{\lambda}_X - \lambda_X) + \frac{\partial \widehat{U}_1^+(\lambda_X, \lambda_Y)}{\partial \lambda_Y}(\widehat{\lambda}_Y - \lambda_Y) + o_p\left(\frac{1}{\sqrt{m}}\right) + o_p\left(\frac{1}{\sqrt{n}}\right)$$

$$= \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{m} \frac{\partial \mathbb{E}\left[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | \tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, \lambda_X, \lambda_Y\right]}{\partial \lambda_X}(\widehat{\lambda}_X - \lambda_X) + \dots$$

[2] under smoothness conditions satisfied for the exponential model

New decomposition

We can use a first order Taylor expansion[2]:

$$\widehat{U}_1^+(\widehat{S}_X, \widehat{S}_Y) - \widehat{U}_1^+(S_X, S_Y) = \widehat{U}_1^+(\widehat{\lambda}_X, \widehat{\lambda}_Y) - \widehat{U}_1^+(\lambda_X, \lambda_Y)$$

$$= \frac{\partial \widehat{U}_1^+(\lambda_X, \lambda_Y)}{\partial \lambda_X}(\widehat{\lambda}_X - \lambda_X) + \frac{\partial \widehat{U}_1^+(\lambda_X, \lambda_Y)}{\partial \lambda_Y}(\widehat{\lambda}_Y - \lambda_Y) + o_p\left(\frac{1}{\sqrt{m}}\right) + o_p\left(\frac{1}{\sqrt{n}}\right)$$

$$= \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{m} \frac{\partial \mathbb{E}\left[\mathbb{1}_{x_{1i} \geq y_{1j} + \tau_1} | \tilde{x}_{1i}, \varepsilon_{1i}, \tilde{y}_{1j}, \eta_{1j}, \lambda_X, \lambda_Y\right]}{\partial \lambda_X}(\widehat{\lambda}_X - \lambda_X) + \dots$$

Similar for Kaplan Meier (KM) but with more complex formulas!

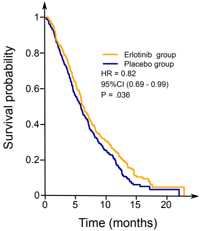
[2] under smoothness conditions satisfied for the exponential model

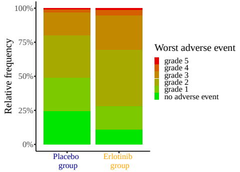
Simulation results (2/2)

- censoring times follow an exponential distribution
- computation time for the standard error: overhead of a factor 1 (n=30) to 18 (n=1000)

Here $\tau = 0.5$ is a large threshold (approx. median survival time)

Back to the motivating example





Priority	$\widehat{\delta}$	$\widehat{\Delta}$	CI($\widehat{\Delta}$)	p-value ($\Delta = 0$)
1 (survival, 2 months)	8.5%	8.5%	[-0.6%;17.5%]	0.066
2 (adverse event)	-5.1%	3.4%	[-6%;12.8%]	0.48

- some evidence for a benefice in survival
- little evidence for a positive benefit-risk balance

Wrapping-up

Net benefit to quantify benefit-risk balance

- $\Delta = \mathbb{P}[X \geq Y + \tau] - \mathbb{P}[X \geq Y + \tau]$
- hierarchy of outcomes, with thresholds of clinical relevance

Estimation (see Ozenne et al. (2021) for details):

- Peron's scoring rule to handle right-censoring
- U-statistic + Taylor expansion to quantify uncertainty
- R package `BuyseTest`

Applications:

- power calculation
- handling measurements with detection limit
- multiple testing adjustment
- handling heteroschedasticity (e.g. permutation test)

Challenges

Reliable inference in small sample / large thresholds:

- `tanh` transformation
- permutation test

Causal interpretation

- not straightforward, see Fay et al. (2018)

Unknown tail of the survival distribution

- add-hoc correction (Péron et al., 2021)
- restricted net benefit

Reference I

Fay, M. P., Brittain, E. H., Shih, J. H., Follmann, D. A., and Gabriel, E. E. (2018). Causal estimands and confidence intervals associated with wilcoxon-mann-whitney tests in randomized experiments. *Statistics in medicine*, 37(20):2923–2937.

Lee, A. J. (1990). *U-statistics: Theory and Practice*. CRC Press.

Moore, M. J., Goldstein, D., Hamm, J., Figer, A., Hecht, J. R., Gallinger, S., Au, H. J., Murawa, P., Walde, D., Wolff, R. A., et al. (2007). Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase iii trial of the national cancer institute of canada clinical trials group. *Journal of clinical oncology*, 25(15):1960–1966.

Reference II

Ozenne, B., Budtz-Jørgensen, E., and Péron, J. (2021). The asymptotic distribution of the net benefit estimator in presence of right-censoring. *Statistical Methods in Medical Research*, 30(11):2399–2412.

Péron, J., Buyse, M., Ozenne, B., Roche, L., and Roy, P. (2018). An extension of generalized pairwise comparisons for prioritized outcomes in the presence of censoring. *Statistical methods in medical research*, 27(4):1230–1239.

Péron, J., Idlhaj, M., Maucort-Boulch, D., Giai, J., Roy, P., Collette, L., Buyse, M., and Ozenne, B. (2021). Correcting the bias of the net benefit estimator due to right-censored observations. *Biometrical Journal*, 63(4):893–906.

Reference III

Randles, R. H. (1982). On the asymptotic normality of statistics with estimated parameters. *The Annals of Statistics*, pages 462–474.

Zhang, S., LeBlanc, M. L., and Zhao, Y.-Q. (2022). Restricted survival benefit with right-censored data. *Biometrical Journal*, 64(4):696–713.

Peron scoring rule

(ε_i,η_j)	$\tilde{x}_i - \tilde{y}_j > \tau$	$\tilde{x}_i - \tilde{y}_j < -\tau$	$\lvert\tilde{x}_i - \tilde{y}_j\rvert < \tau$
(1, 1)	1	-1	0
(0, 1)	1	$\frac{S_X(\tilde{y}_j+\tau)+S_X(\tilde{y}_j-\tau)}{S_X(\tilde{x}_i)} - 1$	$\frac{S_X(\tilde{y}_j+\tau)}{S_X(\tilde{x}_i)}$
(1, 0)	$1 - \frac{S_Y(\tilde{x}_i+\tau)+S_Y(\tilde{x}_i-\tau)}{S_Y(\tilde{y}_j)}$	-1	$-\frac{S_Y(\tilde{x}_i+\tau)}{S_Y(\tilde{y}_j)}$
(0, 0)	A	B	C

$$A = 1 - \frac{S_Y(\tilde{x}_i-\tau)}{S_Y(\tilde{y}_j)} - \frac{\int_{\tilde{x}_i-\tau}^{\infty} S_X(t+\tau)dS_Y(t) + \int_{\tilde{x}_i}^{\infty} S_Y(t+\tau)dS_X(t)}{S_X(\tilde{x}_i)S_Y(\tilde{y}_j)}$$

$$B = -1 + \frac{S_X(\tilde{y}_j-\tau)}{S_X(\tilde{x}_i)} + \frac{\int_{\tilde{y}_j-\tau}^{\infty} S_Y(t+\tau)dS_X(t) + \int_{\tilde{y}_j}^{\infty} S_X(t+\tau)dS_Y(t)}{S_X(\tilde{x}_i)S_Y(\tilde{y}_j)}$$

$$C = \frac{-\int_{\tilde{y}_j}^{\infty} S_X(t+\tau)dS_Y(t) + \int_{\tilde{x}_i}^{\infty} S_Y(t+\tau)dS_X(t)}{S_X(\tilde{x}_i)S_Y(\tilde{y}_j)}$$

Trivial example of H-decomposition

The estimator of the variance:

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 = \frac{1}{\binom{n}{2}}\sum_{i=1}^{n}\sum_{i<j}\frac{(x_i - x_j)^2}{2}$$

is a U-statistic of order 2 with kernel $h(x_1, x_2) = \frac{(x_1 - x_2)^2}{2}$.
Its H-decomposition of $\hat{\sigma}^2$ is:

$$\hat{\sigma}^2 = \sigma^2 + \frac{1}{n}\sum_{i=1}^{n}\left((x_i - \mu)^2 - \sigma^2\right) - \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{j<i}(X_i - \mu)(X_j - \mu)$$

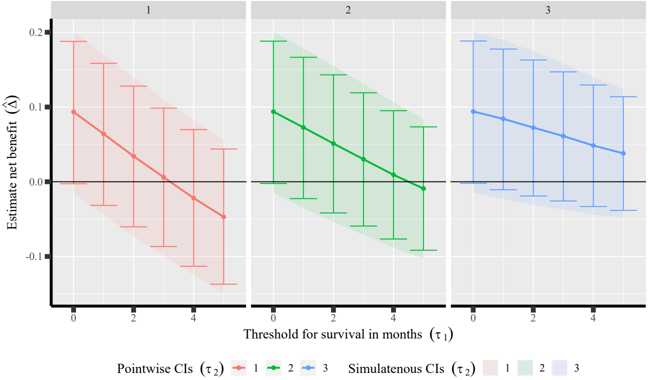
It is therefore asymptotically normally distributed with variance:

$$\mathbb{V}ar\left[\hat{\sigma}^2\right] \xrightarrow[n\to\infty]{} \mathbb{V}ar\left[\frac{1}{n}\sum_{i=1}^{n}\left((x_i - \mu)^2 - \sigma^2\right)\right]$$

$$\xrightarrow[n\to\infty]{} \frac{\mathbb{E}\left[(x_i - \mu)^4\right] - \left(\sigma^2\right)^2}{n}$$

Example of sensitivity analysis

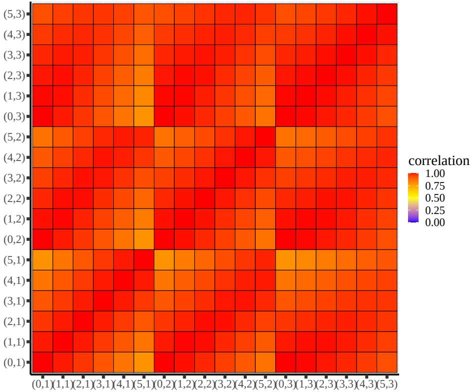
Repeating the analysis varying the thresholds:



Correlation plot of the sensitivity analysis

Jointly normally distributed estimates (asymptotically)

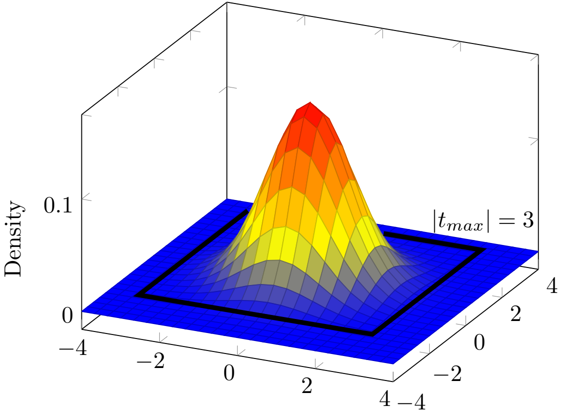
- correlation of 0.79 or above



Adjustment for multiple comparisons

Simulatenous confidence intervals:

- max-test adjustment instead of Bonferroni leveraging the high correlation



Causal interpretation

Ideal causal parameter:

$$\psi = \mathbb{P}[X_i > Y_i]$$

Mann-Whitney parameter:

$$\phi = \mathbb{P}[X_i > Y_j]$$

Causal parameter associated with the Mann-Whitney parameter:

$$\widetilde{\psi} = \mathbb{E}[G(X_i) - G(Y_i)] + 0.5$$

with $G(z) = \frac{\mathbb{P}[X<z]+\mathbb{P}[Y<z]}{2}$

"The value $G(X_i)$ represents where the i-th subject's treatment response falls, in terms of quantiles, among all potential responses, under either arm, in the population.

- For example, $G(X_i) = 0.90$ means that its response when on treatment are about as good or better than 90% of all responses."