

従属打ち切りの下での治療効果推定のための 要因計画・要因分析 —コピュラに基づく方法—

江村 剛志¹・室谷 健太²

(受付 2024 年 10 月 15 日；改訂 2025 年 2 月 11 日；採択 2 月 12 日)

要 旨

治療などの要因が生存時間に与える効果を推定・検定するとき、古典的な正規分布に基づく推測法は適切でない。生存時間の解析では、分布型に依存しないノンパラメトリック法が近年注目を集めている。本論文では、まず独立打ち切りの仮定下での生存時間の要因計画・要因分析のためのノンパラメトリック法を紹介し、それをコピュラモデルに基づいた従属打ち切りの仮定下での解析に拡張した方法を紹介する。続いて、この手法を R 関数 `surv.factorial()` を用いて実行する方法を解説する。これまで調べられていなかった、大域的検定(オムニバス検定)と局所的検定の性能を調べるシミュレーションを行い、これらの検出力特性を明らかにする。最後に、コピュラモデルに基づく手法を、結腸がんの切除手術を受けた被験者のデータに適用し、結果を吟味する。データ解析のための R コードをウェブ上の Supplement で提供する。

キーワード：アルキメデスコピュラ、カプラン・マイヤー推定量、コピュラ・グラフィック推定量、実験計画法、マン・ホイットニー検定、マン・ホイットニー効果。

1. はじめに

実験計画法における要因計画法・要因分析法は、治療などの要因(Factor)が個体に与える効果を推定・検定するための手法である(Montgomery, 2019)。単純に 2 つの治療法を比較する無作為化比較試験に加え、一元配置実験、二元配置実験など、様々な要因の複雑な効果を評価するための統一的なフレームワークが実験計画法に集約されている。例えば、3 つの治療を無作為に割り付ける無作為化比較試験では、全ての治療が等しい効果という大域的仮説の検定に加え、ある 2 つの治療のペアを比較する検定も考えうる。また、治療と性別で 2 つの要因の効果を調べる二元配置実験は、治療の主効果に加え、治療と性別の交互作用効果を推定・検定する(松井, 2023)。

要因計画・要因分析でよく知られている手法は古典的な分散分析(Analysis of Variance, ANOVA)法である(東京大学教養学部統計学教室, 1992)。これは、個体の応答変数が正規分布に従う場合に、治療効果の推定・検定を行う手法である。分散分析において最も基礎的な検定法は、全ての治療群の応答(の平均値)が等しいという大域的な帰無仮説を検証する F -検定で

¹ 広島大学 情報科学部：〒739-8511 広島県東広島市鏡山一丁目 3 番 2 号; takeshiemura@gmail.com

² 久留米大学 医学部：〒830-0011 福岡県久留米市旭町 67; kmurotani@med.kurume-u.ac.jp

ある。この要因全体に効果が無いことを示すための大域的な検定は“オムニバス検定”とも呼ばれる (Mishra et al., 2019; Futschik et al., 2019)。オムニバス検定の棄却は、考慮している治療群間に何らかの差が内在していることを示唆するため、その後に2群間のペアワイズ比較や、主効果や交互作用などの詳細な推定・検定を行う根拠となる。

がんの臨床試験では、通常は2つの治療の無作為化比較試験が行われるが、3つの治療の優劣を比較する試験が行われることもある (Moertel et al., 1995)。また、がんの臨床試験における治療効果の検証において、2要因×2水準 (= 4つの治療)の要因計画・要因分析が行われることがある。例えば、進行子宮頸がん患者の生存に対する治療の効果を調べた試験 (Tewari et al., 2014) や、転移性前立腺がん患者を対象とした進行中の第III相臨床試験 (Fizazi et al., 2022) を参照されたい。2017年の調査では、第III相がん治療の2×2実験計画の臨床試験が30件特定された (Freidlin and Korn, 2017)。

がん臨床試験のように、応答変数が被験者の生存時間の場合、応答変数の正規性に基づく平均・分散を基にした古典的な手法の利用は適切でない。生存時間データは右に歪んだ形状を持つ場合も多く、正規分布に従わないだけでなく、平均・分散は存在しないこともある (武富・山本, 2023)。このため、生存時間解析における多群の治療効果比較における要因分析では、分布型に依存しないノンパラメトリックな方法が近年注目を集めている。例えば、マン・ホイットニー効果 (Dobler and Pauly, 2020)、標本分割を用いたペアワイズ効果 (Gorfine et al., 2020)、生存期間中央値に対する効果 (Ditzhaus et al., 2021)、累積ハザードに対する効果 (Ditzhaus et al., 2023) などで治療効果を定義し、ノンパラメトリック推定・検定を行う。

生存時間解析で難しい点は“打ち切り”の扱いである (Klein and Moeschberger, 2003)。上述のノンパラメトリック要因計画・要因分析法で提案されている生存時間の統計手法は、打ち切り発生が生存リスクに影響しないという“独立打ち切り”の仮定の基に構築されている。とりわけこれら手法は、独立打ち切りの仮定を基に導出された Kaplan-Meier (KM) 推定量を適用しているものもある。しかしながら、KM 推定量は独立打ち切りが成立しないとき、生存関数の推定にバイアスをもたらす (Rivest and Wells, 2001; Emura and Hsu, 2020; 江村, 2023; 斎藤・室谷, 2024)。

独立でない打ち切り (従属打ち切りと呼ばれる) は、被験者追跡を要する医学データの分析で現れやすい (Andersen and Perme, 2010; Staplin et al., 2015; Xu et al., 2018; Moradian et al., 2019; Collett, 2023; Li et al., 2023; Das et al., 2024)。典型的な例は、健康状態の悪化による被験者の脱落 (ドロップアウト) による打ち切りである。例えば、Staplin et al. (2015) によると、追跡の打ち切りは被験者が肝移植の待機リストから削除されたときに発生する。この例において、死期が近い被験者は移植を受ける可能性が高くなるため、“従属打ち切り”をもたらす可能性がある。より一般に、病状悪化のために追跡不能となった被験者がその後に死亡している疑いのあるデータ (6節を参照) では、死亡時間と打ち切り時間が正相関している可能性がある。

独立打ち切りの仮定に立脚しない代替手法の開発も数多く行われている (Lu and Zhang, 2012; Staplin et al., 2015; Dettoni et al., 2020; Deresa and Van Keilegom, 2020, 2025; Deresa and Van Keilegom, 2020; Schneider et al., 2020, 2022)。例えば、従属打ち切り下での KM 推定量はバイアスを持つが、Andersen and Perme (2010) は、観測可能な共変量の情報を用いて、バイアスを削減する手法を提案している。Emura and Chen (2016) は、従属打ち切りと生存時間の従属構造をクレイトンコピュラでモデル化し、単変量 Cox 回帰法による変数選択の代替手法を提案している。同様に、Emura and Hsu (2020) もコピュラモデルを用いて、従属打ち切り下でのバイアスを修正したマン・ホイットニー効果量に基づく推定・検定法を提案している。他にも、従属打ち切り問題への対処のための様々な手法は、Emura and Chen (2018) と Collett (2023) の書籍にまとめられているが、その後多くの手法が提案されている (Jo et al., 2023; Gares et

al., 2024; Gottselig et al., 2025; Delhelle and Van Keilegom, 2025). ただし, 要因計画・要因分析における従属打ち切りの問題への対処法を提案した手法は Emura et al. (2024) に限られる.

本論文では, まず Dobler and Pauly (2020) が提案した生存時間解析のための要因計画・要因分析手法を紹介する(2節). その後, コピュラモデルに基づき従属打ち切りを考慮することができる要因分析法 (Emura et al., 2024) を詳説する(3節). この手法では, 打ち切り時間と生存時間の従属構造を想定した感度分析を行う手法である. 4節では, これら解析を R 関数 `surv.factorial()` を用いて実行する方法を解説する. 5節ではシミュレーションを行い, これまで調べられていなかった, 大域的検定(オムニバス検定)と局所的検定の性能を調べるシミュレーションを行い, これらの検出力特性を明らかにする. また本論文では, 新たなデータ解析例として, 結腸がんの切除手術を受けた被験者のデータへの適用例を紹介する(6節). 7節を本論文の結びとする. 尚, データ解析に使用した R コードはオンライン Supplement から入手可能である.

2. 生存時間の要因計画・要因分析

本節では, 生存時間データに基づく要因計画・要因分析の手法 (Dobler and Pauly, 2020) の概要を紹介する. まず, 一元配置, 二元配置, その他の様々な実験計画において統一的に定義可能な治療効果を定式化する. その後, 推定・検定法を紹介する.

2.1 治療効果

治療などの要因が生存時間に与える影響を“治療効果”と呼ぶ. ただし, 要因は性別や環境要因などの実験的に制御出来ないものでも良い. 生存時間のデータは正規分布に従うことが稀であるため, 平均の等価性を検定するための古典的な分散分析の利用は適切でない. 故に, 治療効果の定式化には, ノンパラメトリックな治療効果指標を要する. 以下で紹介する手法は, 2群のノンパラメトリック検定で良く知られている「マン・ホイットニー検定」の考え方を拡張した治療効果の推定・検定法 (Dobler and Pauly, 2020; Brunner et al., 2017; Dobler et al., 2020) である.

治療が d 通りあるとし, 各治療下での生存時間を T_i , その生存関数を $S_i(t) = P(T_i > t)$, $i = 1, 2, \dots, d$, とする. i 番目と ℓ 番目を比較した“ペアワイズ効果”は, 次のようになる.

$$(2.1) \quad w_{i\ell} = P(T_i > T_\ell) + \frac{1}{2}P(T_i = T_\ell) = - \int S_i^\pm(t)dS_\ell(t),$$

ここで $S_i^\pm(t) = \{S_i(t+) + S_i(t-)\}/2$, $S_i(t+) = \lim_{s \downarrow t} S_i(s)$, $S_i(t-) = \lim_{s \uparrow t} S_i(s)$ である. この $w_{i\ell}$ は, i 番目の治療を受けた被験者が, ℓ 番目の治療を受けた被験者より長く生存する確率である. もし $w_{i\ell} > 1/2$ であれば, i 番目の治療が ℓ 番目の治療に比べて有益であることを示す(もし $w_{i\ell} < 1/2$ であれば, その逆である). なお, $w_{ii} = 1/2$ である.

2標本($d = 2$)の場合, ペアワイズ効果 w_{12} のみを用いて推定・検定を行えば良い (Efron, 1967; Koziol and Jia, 2009; Dobler and Pauly, 2018; Emura and Hsu, 2020; Nowak et al., 2022) が, 本論文で注目するのは主に $d \geq 3$ の場合である. このように 3種類以上の治療法がある場合, ペアワイズ効果ではなく, 一般的な治療と対比した“相対的な効果”で治療効果を定式化すると良い. この定式化は, 古典的な分散分析における治療効果を, 群平均と一般平均の差で定義していることと同じ考え方である. 以下でその方法を記述する.

ノンパラメトリック分散分析 (Brunner and Puri, 2001; Brunner et al., 2017) の枠組みを用い, Dobler and Pauly (2020) は, i 番目の治療の相対的な治療効果を次式で定義した.

$$p_i = \frac{1}{d} \sum_{\ell=1}^d w_{i\ell} = - \int S_i^\pm(t) d\bar{S}(t) = P(T_i > \bar{T}) + \frac{1}{2} P(T_i = \bar{T}),$$

ここに $\bar{S} = \sum_{\ell=1}^d S_\ell/d$ と定義し, これは集団全体での生存関数である. この p_i は, i 番目の治療を受けた被験者が, 被験者全体の集団より長く生存する割合と解釈できる. 治療が d 通りあるため, これらの治療効果をベクトルでまとめると次のようになる:

$$(2.2) \quad \mathbf{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_d \end{pmatrix} = \mathbf{A}\mathbf{w}, \quad \mathbf{w} = \left(w_{11}, \dots, w_{1d} : w_{21}, \dots, w_{2d} : \dots : w_{d1}, \dots, w_{dd} \right)',$$

ここで $\mathbf{A} = \mathbf{I}_d \otimes (\mathbf{1}'_d/d)$ であり, \mathbf{I}_d は $d \times d$ 単位行列, $\mathbf{1}'_d = (1, \dots, 1)$ は長さ d のベクトル, \otimes はクロネッカー積である. クロネッカー積の定義と式(2.2)の証明は付録 1 を参照されたい.

治療効果 \mathbf{p} に関する帰無仮説・対立仮説は, 対比行列 \mathbf{C} を用いて

$$H_0: \mathbf{C}\mathbf{p} = \mathbf{0} \quad \text{vs.} \quad H_1: \mathbf{C}\mathbf{p} \neq \mathbf{0}$$

と表せる. 最も単純な仮説検定の例は, $\mathbf{C} = \mathbf{P}_d = \mathbf{I}_d - \mathbf{1}_d\mathbf{1}'_d/d$ で定まる大域的な検定問題

$$H_0: p_1 = p_2 = \dots = p_d \quad \text{vs.} \quad H_1: p_i \neq p_\ell, \quad \exists(i, \ell).$$

である(付録 1). 他の検定問題は, 例えば局所仮説 $H_0: p_1 = p_d$ は $\mathbf{C} = (1, 0, \dots, 0, -1)$ など, 対比行列を適宜指定することによって定式化できる.

少し複雑な例として, 二元配置を考える. a 個の水準を持つ“因子 A”と b 個の水準を持つ“因子 B”を考える. 因子計画では, 全ての治療を組み合わせた $a \times b$ 個の治療効果を考えるため,

$$\mathbf{p} = \left(p_{11}, \dots, p_{1b} : p_{21}, \dots, p_{2b} : \dots : p_{a1}, \dots, p_{ab} \right)',$$

と書け, ここに p_{jk} は因子 A の j 番目の水準と因子 B の k 番目の水準での治療効果である. 例えば, 因子 A の主効果の有無を検定する場合は $\mathbf{C} = \mathbf{P}_a \otimes \mathbf{1}'_b/b$ を指定すると,

$$H_0: p_{1\cdot} = p_{2\cdot} = \dots = p_{a\cdot} \quad \text{vs.} \quad H_1: p_{i\cdot} \neq p_{\ell\cdot} \text{ for } \exists(i, \ell), i \neq \ell,$$

となり, ここに $p_{i\cdot} = \sum_{\ell=1}^b p_{i\ell}$ である. 同様に, 因子 B の主効果の有無の検定は $\mathbf{C} = \mathbf{1}'_a/a \otimes \mathbf{P}_b$, 交互作用の有無の検定 $\mathbf{C} = \mathbf{P}_a \otimes \mathbf{P}_b$ と指定すれば良い.

2.2 独立打ち切り下での要因分析法

Dobler and Pauly (2020) は, 生存時間データに基づき, 治療効果の推定・検定を行うノンパラメトリックな手法を提案した. この手法では, 独立打ち切りの仮定の下, 治療効果の推定量を

$$\hat{\mathbf{p}} = \begin{pmatrix} \hat{p}_1 \\ \vdots \\ \hat{p}_d \end{pmatrix} = \mathbf{A}\hat{\mathbf{w}}, \quad \hat{\mathbf{w}} = \left(\hat{w}_{11}, \dots, \hat{w}_{1d} : \hat{w}_{21}, \dots, \hat{w}_{2d} : \dots : \hat{w}_{d1}, \dots, \hat{w}_{dd} \right)'$$

としており, ここにペアワイズ効果の推定量 $\hat{\mathbf{w}}$ は生存関数 $S_i(\cdot)$ の KM 推定量 $\hat{S}_i(\cdot)$ を用い,

$$\hat{w}_{i\ell} = - \int \hat{S}_i^\pm(t) d\hat{S}_\ell(t),$$

としており, ここに $\hat{S}_i^\pm(t) = \{\hat{S}_i(t+) + \hat{S}_i(t-)\}/2$ である. ただし, $\hat{S}_i(t)$ が $S_i(t)$ の一致推定量

となる範囲は $t \in [0, \tau]$ である ($\tau > 0$ はある定数) から, 上式の積分に修正が必要となる (3 節で説明する). 全群のサンプルサイズ(被験者数)を N とすると, $N \rightarrow \infty$ のとき漸近正規近似 $\sqrt{N}(\hat{p} - p) \sim \mathcal{N}(\mathbf{0}, V)$ が導出されており, ここに V は共分散行列であり, その推定量 \hat{V} も提案されている. これより \hat{p}_i の標準誤差(SE) および p_i の信頼区間が求まる.

帰無仮説 $H_0: C\mathbf{p} = \mathbf{0}$ からの逸脱は統計量 $C\hat{p}$ と $\mathbf{0}$ の差で調べることが出来る. Dobler and Pauly (2020) は次の F -統計量を提案している:

$$F_N = \frac{N\hat{p}'T\hat{p}}{\text{tr}(T\hat{V})} = \frac{N(C\hat{p})'(CC')^+(C\hat{p})}{\text{tr}[(CC')^+\hat{V}CC']},$$

ここに, $T = C'(CC')^+C$ は射影行列, $\text{tr}(\cdot)$ はトレース, $(CC')^+$ は CC' の一般化逆行列(ムーア・ペンローズ逆行列)である(一般化逆行列の計算は, 例えば R であれば関数 `ginv(.)` を用いれば良い).

統計量 F_N の $N \rightarrow \infty$ のときの漸近帰無分布は, 重みづけカイ二乗分布の和であることが示されており, 漸近帰無分布の平均は 1 である (Dobler and Pauly, 2020). 対立仮説の下で, F_N は無限大に確率収束することも示されるため, 条件 $F_N > c_{N,\alpha}$ で帰無仮説を棄却する検定方式が採用され, ここに棄却限界値 $c_{N,\alpha}$ は数値的に求める (Dobler and Pauly, 2020). この検定は独立打ち切りの仮定の下で妥当であり, 従属打ち切り下では妥当性は無いことに注意する.

2.3 大域的仮説と局所的仮説の関係

Dobler and Pauly (2020) の方法ならびに多くの要因分析では, 多数の群の間のどこに差があるか事前に分からぬいため, 最初に大域的仮説を行うことがある. もし, 大域的仮説が棄却されれば, より詳細な群間差を検出するための局所的仮説を検定して行く. 例えば, 3 群の場合では以下の 3 つの局所的検定問題と大域的検定問題が考えられる:

$$H_0: p_1 = p_2 \quad \text{vs.} \quad H_1: p_1 \neq p_2,$$

$$H_0: p_1 = p_3 \quad \text{vs.} \quad H_1: p_1 \neq p_3,$$

$$H_0: p_2 = p_3 \quad \text{vs.} \quad H_1: p_2 \neq p_3$$

$$H_0: p_1 = p_2 = p_3 \quad \text{vs.} \quad H_1: p_1 \neq p_2, \quad \text{or} \quad p_1 \neq p_3, \quad \text{or} \quad p_2 \neq p_3.$$

通常の手続きとしては, 最後の大域的仮説検定で棄却となった場合に限り, 最初の 3 つの局所的仮説を検定して行く. ここで問題になるのは, 大域的仮説検定と局所的仮説検定の検出力の違いである. 例えば, 大域的仮説が棄却された場合に局所的仮説を検定する手順で進める場合は, 大域的仮説検定の検出力が十分高い必要がある. この検出力の性能評価は生存時間の要因分析の文脈で議論されてこなかったため, 5 節で数値的に議論する.

3. 従属打ち切り下での推定・検定

まず, 従属打ち切り下での推定・検定を行うための生存コピュラモデルを定式化する. 次に, Rivest and Wells (2021) による生存関数の推定量を紹介する. その後, 要因分析のために提案された Emura et al. (2024) による治療効果の推定法・検定法を紹介する.

3.1 生存関数の推定

治療が d 通りあるとし, $i = 1, 2, \dots, d$ は治療群を表すとする. i 番目の治療に n_i 人の被験者が割り付けられるとし, $j = 1, \dots, n_i$ は被験者番号を表すとする. 各被験者の生存時間を T_{ij} , 打ち切り時間を U_{ij} とすると, 右打ち切り下で観測されるデータは

$\{(X_{ij}, \delta_{ij}); i = 1, \dots, d, j = 1, \dots, n_i\}$ であり, ここで $X_{ij} = \min(T_{ij}, U_{ij})$, $\delta_{ij} = 1\{T_{ij} \leq U_{ij}\}$, $1\{\cdot\}$ は指示関数である. 全被験者数(サンプルサイズ)は $N = \sum_{i=1}^d n_i$ である.

従属打ち切りの構造を定めるため, 打ち切り時間と生存時間の同時生存関数に次のモデルを仮定する:

$$(3.1) \quad P(T_{ij} > t, U_{ij} > u) = C_{i\theta_i}(S_i(t), G_i(u)), \quad i = 1, 2, \dots, d,$$

ここで $C_{i\theta_i}(\cdot, \cdot)$ はコピュラ (Nelsen, 2006; Durante and Sempi, 2016; 江村, 2025) で, θ_i は従属の正負や強さを表すパラメータ, $S_i(t) = P(T_{ij} > t)$, $G_i(u) = P(U_{ij} > u)$ は周辺生存関数である. クレイトンコピュラ (Clayton, 1978) を仮定する場合,

$$C_{i\theta_i}(u, v) = (u^{-\theta_i} + v^{-\theta_i} - 1)^{-\frac{1}{\theta_i}}, \quad \theta_i > 0, \quad i = 1, 2, \dots, d,$$

とし, ここに θ_i はケンドール順位相関(ケンドールのタウ)に変換すると $\tau_{\theta_i} = \theta_i/(\theta_i + 2)$ となる. 独立コピュラ $C_{i\theta_i}(u, v) = uv$, $i = 1, 2, \dots, d$, は独立打ち切りの仮定に相当する. 他のコピュラの例としては, グンベルコピュラやフランクコピュラがある (杉本・田中, 2023; 江村, 2025). 多くのコピュラでは, 極限 $\theta_i \rightarrow 0$ は $C_{i\theta_i}(u, v) \rightarrow uv$ となるため, モデル(3.1)は独立打ち切りのモデルを包含している. また, モデル(3.1)は, 周辺生存関数 $S_i(\cdot)$, $G_i(\cdot)$ にモデル化は行わず, 相関構造のみをモデル化している.

モデル(3.1)の下で, 観測データ $\{(X_{ij}, \delta_{ij}); i = 1, \dots, d, j = 1, \dots, n_i\}$ から, 周辺生存関数 $S_i(\cdot)$ のノンパラメトリック推定が可能となることが知られている. Rivest and Wells (2001) は, モデル(3.1)がアルキメデスコピュラである場合の, 具体的な推定量を提案している. アルキメデスコピュラとは, 生成素と呼ばれる関数 $\phi_{i\theta_i}$ を用いて次のように表せるモデルである:

$$(3.2) \quad P(T_{ij} > t, U_{ij} > u) = \phi_{i\theta_i}^{-1}[\phi_{i\theta_i}\{S_i(t)\} + \phi_{i\theta_i}\{G_i(u)\}], \quad i = 1, \dots, d,$$

ここで, 生成素 $\phi_{i\theta_i}$ は減少かつ Convex な連続関数であり $\phi_{i\theta_i}(0) = \infty$ かつ $\phi_{i\theta_i}(1) = 0$ を満足する必要がある. 同じ群内でタイが無い ($j \neq k$ に対し $X_{ij} \neq X_{ik}$) と仮定する. このとき, モデル(3.2)の仮定の下, Rivest and Wells (2001) は次の Copula-Graphic(CG) 推定量

$$\hat{S}_i^{\text{CG}}(t) = \phi_{i\theta_i}^{-1} \left[\sum_{j: X_{ij} \leq t, \delta_{ij}=1} \left\{ \phi_{i\theta_i} \left(\frac{\bar{Y}_i(X_{ij}) - 1}{n_i} \right) - \phi_{i\theta_i} \left(\frac{\bar{Y}_i(X_{ij})}{n_i} \right) \right\} \right], \quad 0 \leq t \leq \max_j(X_{ij}),$$

を提案しており, ここに $\bar{Y}_i(x) = \sum_{j=1}^{n_i} 1\{X_{ij} \geq x\}$ である. モデル(3.2)が正しく指定されている場合, ある正則条件下で推定量は $\hat{S}_i^{\text{CG}}(t)$ は $S_i(t)$ の一致推定量となる (Rivest and Wells, 2001).

クレイトンコピュラの下では, 生成素 $\phi_{i\theta_i}(t) = (t^{-\theta_i} - 1)/\theta_i$, $\theta_i > 0$, を用い,

$$\hat{S}_i^{\text{CG}}(t) = \left[1 + \sum_{j: X_{ij} \leq t, \delta_{ij}=1} \left\{ \left(\frac{\bar{Y}_i(X_{ij}) - 1}{n_i} \right)^{-\theta_i} - \left(\frac{\bar{Y}_i(X_{ij})}{n_i} \right)^{-\theta_i} \right\} \right]^{-\frac{1}{\theta_i}}$$

となる. グンベルコピュラの下では, 生成素 $\phi_{i\theta_i}(t) = (-\log t)^{\theta_i+1}$, $\theta_i \geq 0$, を用い

$$\hat{S}_i^{\text{CG}}(t) = \exp \left[- \left(\sum_{j: X_{ij} \leq t, \delta_{ij}=1} \left[\left\{ -\log \left(\frac{\bar{Y}_i(X_{ij}) - 1}{n_i} \right) \right\}^{\theta_i+1} - \left\{ -\log \left(\frac{\bar{Y}_i(X_{ij})}{n_i} \right) \right\}^{\theta_i+1} \right] \right)^{\frac{1}{\theta_i+1}} \right].$$

さらに, フランクコピュラの下で生成素 $\phi_{i\theta_i}(t) = -\log[(e^{-\theta_i t} - 1)/(e^{-\theta_i} - 1)]$ を適用すると,

$$\hat{S}_i^{\text{CG}}(t) = -\frac{1}{\theta_i} \log \left[1 + (e^{-\theta_i} - 1) \exp \left(\sum_{j: X_{ij} \leq t, \delta_{ij}=1} \log \left\{ \frac{\exp \left(-\theta_i \frac{\bar{Y}_i(X_{ij})-1}{n_i} \right) - 1}{\exp \left(-\theta_i \frac{\bar{Y}_i(X_{ij})}{n_i} \right) - 1} \right\} \right) \right]$$

となる。加えて、独立コピュラの下では $\phi_{i\theta_i}(t) = -\log(t)$ を用い、次の KM 推定量となる

$$\hat{S}_i^{\text{CG}}(t) = \hat{S}_i(t) = \prod_{j: X_{ij} \leq t, \delta_{ij}=1} \left[1 - \frac{1}{\bar{Y}_i(X_{ij})} \right], \quad 0 \leq t \leq \max_j(X_{ij}).$$

CG 推定量を用いると、さまざまな従属打ち切りの相関構造を想定した下での生存関数の推定および感度分析が可能となる。もし $\phi_{i\theta_i}$ が母集団の相関構造を正しく特定していれば、CG 推定量は一致推定となるが、誤特定している場合はバイアスが生じる。Rivest and Wells (2001) の感度分析によると、母集団のケンドール順位相関の大きさを θ_i が捉えている場合、 $\phi_{i\theta_i}$ の関数型に誤特定があっても CG 推定量のバイアスは KM 推定量のそれよりも小さくなると数値的に示している。

上述の CG 推定量を計算するには、R パッケージ *compound.Cox* (Emura et al., 2019) 内にある関数 *CG.Clayton()*, *CG.Frank()*, *CG.Gumbel()* を用いれば良い。

3.2 治療効果の推定

従属打ち切り下で生存関数の推定が行えることを説明したが、要因分析において推定・検定の対象としているのは治療効果 p である。前述の CG 推定量を用いると、独立打ち切り下での推定量 \hat{p} (2.2 節で定義した) は従属打ち切りの下に自然に拡張される。

まず、式(2.1)のペアワイズ効果 $w_{i\ell}$ を次のように推定する：

$$\hat{w}_{i\ell}^{\text{CG}} = - \int \hat{S}_i^{\text{CG}\pm}(t) d\hat{S}_\ell^{\text{CG}}(t).$$

次に、式(2.2)の $p = Aw$ を用い、 p の推定量を次式で定義する (Emura et al., 2024)。

$$\hat{p}^{\text{CG}} = \begin{pmatrix} \hat{p}_1^{\text{CG}} \\ \vdots \\ \hat{p}_d^{\text{CG}} \end{pmatrix} = A\hat{w}^{\text{CG}} = A\left(\hat{w}_{11}^{\text{CG}}, \dots, \hat{w}_{1d}^{\text{CG}} : \hat{w}_{21}^{\text{CG}}, \dots, \hat{w}_{2d}^{\text{CG}} : \dots : \hat{w}_{d1}^{\text{CG}}, \dots, \hat{w}_{dd}^{\text{CG}}\right)'$$

これは、生存関数に分布の仮定を課さないノンパラメトリックな推定量である。

ある正則条件(付録 2)の下、モデル(3.2)が正しければ推定量は漸近正規性を満足するため、 $N \rightarrow \infty$ としたとき近似 $\sqrt{N}(\hat{p}^{\text{CG}} - p) \sim N_d(\mathbf{0}, \mathbf{V})$ が可能となり、ここに \mathbf{V} は共分散行列である。未知の \mathbf{V} はジャックナイフ推定量を N 倍してスケーリングした $\hat{\mathbf{V}} = N \times \mathbf{V}(\hat{p}^{\text{CG}})$ で推定可能であり (Emura et al., 2024)，ここにジャックナイフ推定量は

$$\mathbf{V}(\hat{p}^{\text{CG}}) = \begin{pmatrix} \hat{\sigma}_1^2 & \dots & \hat{\sigma}_{d1} \\ \vdots & \ddots & \vdots \\ \hat{\sigma}_{d1} & \dots & \hat{\sigma}_d^2 \end{pmatrix} = \frac{N-1}{N} \sum_{i=1}^d \sum_{j=1}^{n_i} (\hat{p}^{\text{CG}(-ij)} - \hat{p}^{\text{CG}(.)})(\hat{p}^{\text{CG}(-ij)} - \hat{p}^{\text{CG}(.)})',$$

ここで $\hat{p}^{\text{CG}(-ij)}$ は i 番目の群の j 番目の被験者を除いて計算した推定量、 $\hat{p}^{\text{CG}(.)} = \sum_{i=1}^d \sum_{j=1}^{n_i} \hat{p}^{\text{CG}(-ij)}/N$ はそれらの平均である。原論文の Emura et al. (2024) ではジャックナイフ分散推定量の式の $\frac{N-1}{N}$ の部分を $\frac{N}{N-1}$ としたが、これはタイポであり、本稿では $\frac{N-1}{N}$ と修正した。R 関数 *surv.factorial()* 上での計算プログラムに誤りはなく、修正の必要は無い。

CG 推定量はタイを許容しない都合上、通常のブートストラップ法のような復元抽出による

方法は適用が難しい。また、標準誤差(SE)は $SE(\hat{p}_i^{CG}) = \hat{\sigma}_i$ となり、 p_i の $(1 - \alpha) \times 100\%$ 信頼区間(CI)は

$$[\hat{p}_i^{CG} - z_{\alpha/2} \times SE(\hat{p}_i^{CG}), \quad \hat{p}_i^{CG} + z_{\alpha/2} \times SE(\hat{p}_i^{CG})],$$

ここに $z_{\alpha/2}$ は標準正規分布の上側 $(\alpha/2) \times 100\%$ 点である。信頼区間は上限・下限が p_i の範囲 $[0, 1]$ を超える可能性があるが、その場合必要に応じて、logit 変換などで区間 $[0, 1]$ に収まるようにも良い (Brunner et al., 2017; Perme and Manevski, 2019) が、本論文では特に用いない。

以上の手法において、独立コピュラを適用すると Dobler and Pauly (2020) の手法に帰着する。しかしながら、彼らは、ジャックナイフ法とは異なり、Multiplier ブートストラップ法で SE と信頼区間を構成している。Multiplier ブートストラップ法は、独立打ち切り下では有用とみられるが、上記のコピュラモデル下での適用は難しい。

3.3 治療効果の検定

ある対比行列 C に対する検定問題 $H_0: C\mathbf{p} = \mathbf{0}$ vs. $H_1: C\mathbf{p} \neq \mathbf{0}$ を定式化する。独立打ち切りの仮定下で定義した F_N の式(2.2 節で定義した)に \hat{p}^{CG} を適用すると、次の統計量を得る：

$$F_N^{CG} = \frac{N(\hat{p}^{CG})' \mathbf{T} \hat{p}^{CG}}{\text{tr}(\mathbf{T} \hat{V}^{CG})}$$

ここで、 $\mathbf{T} = \mathbf{C}'(\mathbf{C}\mathbf{C}')^+ \mathbf{C}$ は射影行列、 $\text{tr}(\cdot)$ はトレース、 $\hat{V}^{CG} = N \times \mathbf{V}(\hat{p}^{CG})$ はジャックナイフ推定量を N 倍して共分散行列 \mathbf{V} の推定量にスケーリングしたものである。

統計量 F_N^{CG} の漸近分布について次のような結果が得られている。いま、 $\mathbf{T}\mathbf{V}$ の固有値を $\lambda_i, i = 1, \dots, d$ とすると、 $\sum_{i=1}^d \lambda_i = \text{tr}(\mathbf{T}\mathbf{V})$ である。帰無仮説 H_0 の下で $N \rightarrow \infty$ としたときの F_N^{CG} 漸近分布は、重み付きカイ二乗分布の和

$$F_N^{CG} \sim \frac{\sum_{i=1}^d \lambda_i \chi_i^2(1)}{\sum_{i=1}^d \lambda_i} = \frac{\sum_{i=1}^d \lambda_i \chi_i^2(1)}{\text{tr}(\mathbf{T}\mathbf{V})},$$

であり、ここに $\chi_i^2(1)$ は自由度 1 のカイ二乗分布を表す (Emura et al., 2024)。この重みづけカイ二乗分布の和の上側 $\alpha \times 100\%$ 点を $c_{N,\alpha}$ とする。帰無仮説の下で漸近分布の期待値は、上式から明らかに 1 である。また、対立仮説の下で F_N^{CG} は無限大に確率収束する。よって、漸近的に有意水準 α である検定は棄却領域を $F_N^{CG} > c_{N,\alpha}$ とすれば良い。

この棄却限界 $c_{N,\alpha}$ は未知の行列 \mathbf{V} に依存するため、データから推定する必要がある。次に示すよう、推定法として「シミュレーション法」と「解析的方法」の 2 つが提案されているが、どちらを用いてもほぼ同じような数値結果が得られる(5 節、6 節)。

シミュレーション法による $c_{N,\alpha}$ の推定

打ち切りの無い場合に Brunner et al. (2017) で提案された方法を適用する。まず λ_i の推定値 $\hat{\lambda}_i$ を $\mathbf{T} \hat{V}^{CG}$ の固有値で定める。次に自由度 1 のカイ二乗分布に従う乱数 $\chi_{i,(r)}^2(1)$, $i = 1, \dots, d$, $r = 1, \dots, R$ を生成し、ここに R は大きな数 $R = 1000$ などとする。最後に、

$$F_{N,(r)}^{CG} = \frac{\sum_{i=1}^d \hat{\lambda}_i \chi_{i,(r)}^2(1)}{\text{tr}(\mathbf{T} \hat{V}^{CG})}, \quad r = 1, \dots, R$$

のヒストグラムの上側 $\alpha \times 100\%$ 点として $c_{N,\alpha}$ を推定する。検定の P 値は、ヒストグラムの上側確率 $\sum_{r=1}^R 1\{F_{N,(r)}^{CG} > F_N^{CG}\}/R$ で定める。

解析的方法による $c_{N,\alpha}$ の推定

打ち切りの無い場合に Brunner and Puri (2001) と Brunner et al. (2017) で提案された方法を

適用する。これは重み付きカイ二乗分布の和(漸近分布)をカイ二乗分布で近似する手法である。ある正の定数 f と g に対し、以下の近似

$$\frac{\sum_{i=1}^d \lambda_i \chi_i^2(1)}{\text{tr}(\mathbf{TV})} \approx g \chi^2(f)$$

を考える。ここで、 f と g は上式の右辺と左辺の期待値と分散が同じになるよう定めると、

$$\frac{\sum_{i=1}^d \lambda_i}{\text{tr}(\mathbf{TV})} = gf, \quad 2 \frac{\sum_{i=1}^d \lambda_i^2}{[\text{tr}(\mathbf{TV})]^2} = 2g^2f$$

である。最初の方程式左辺は 1 であるから、 $1 = gf$ となる。これと 2 番目の方程式より f と g が求まり、さらに未知の \mathbf{V} に推定値を代入すると以下となる。

$$c_{N,\alpha} = \frac{\chi^2(1 - \alpha, \hat{f})}{\hat{f}}, \quad \hat{f} = \frac{[\sum_{i=1}^d \hat{\lambda}_i]^2}{\sum_{i=1}^d \hat{\lambda}_i^2} = \frac{[\text{tr}(\hat{\mathbf{TV}}^{\text{CG}})]^2}{\text{tr}(\hat{\mathbf{TV}}^{\text{CG}} \hat{\mathbf{TV}}^{\text{CG}})}.$$

ここで、 $\chi^2(1 - \alpha, \hat{f})$ は自由度 \hat{f} のカイ二乗分布の上側 $100 \times \alpha\%$ 点である。検定の P 値は、カイ二乗分布の上側確率 $1 - F_{\chi^2(\hat{f})}(\hat{f} F_N^{\text{CG}})$ で定め、ここに $F_{\chi^2(\hat{f})}$ は $\chi^2(\hat{f})$ の累積分布関数である。

3.4 フォローアップ期間の制約

定まったフォローアップ期間内で収集されたデータを基にした生存関数の推定は不完全であるから、その結果として治療効果 p の推定も不完全となる。ある上限 $\tau > 0$ を超えた生存時間が確実に打ち切られる状況下において、生存関数 $S_i(t)$ が推定可能な範囲は $t \in [0, \tau]$ に制約され、ここに

$$\tau < \min_i [\sup_u \{u: G_i(u) S_i(u) > 0\}]$$

である (Emura et al., 2024)。これより、生存関数 $S_i(t)$ は $t \in [0, \tau]$ においてのみ推定可能となる。以下、治療効果 p を推定可能な生存関数 $S_i(t)$ の範囲 $t \in [0, \tau]$ を用いたものに修正する。

まずペアワイズ効果を次のように修正する。

$$\begin{aligned} w_{i\ell} &= P[\min(T_i, \tau) > \min(T_\ell, \tau)] + \frac{1}{2} P[\min(T_i, \tau) = \min(T_\ell, \tau)] \\ &= - \int_0^\tau S_i^\pm(t) dS_\ell(t) + \frac{S_i(\tau) S_\ell(\tau)}{2}. \end{aligned}$$

右辺の最後の項は、積分を切斷したことによるペアワイズ効果の減少量を補正している。なお、 $w_{ii} = 1/2$ である。これより、修正したペアワイズ効果 $w_{i\ell}$ の推定量は次のようになる

$$\hat{w}_{i\ell}^{\text{CG}} = - \int_0^\tau \hat{S}_i^{\text{CG}\pm}(t) d\hat{S}_\ell^{\text{CG}}(t) + \frac{\hat{S}_i^{\text{CG}}(\tau) \hat{S}_\ell^{\text{CG}}(\tau)}{2}.$$

これより、治療効果 $\hat{p}^{\text{CG}} = \mathbf{A} \hat{\mathbf{w}}^{\text{CG}}$ が求まる。これと同様に検定法も修正すれば良い。

3.5 感度分析

治療効果の推定・検定を実行するためには、コンピュータの型とパラメータ θ_i を指定する必要がある。しかしながら、従属打ち切りのモデルが同定困難 (Tsiatis, 1975) であるため、CG 推定量に用いるコンピュータもデータから推定することは困難である。このため、コンピュータとそのパラメータの候補をいくつか選択し、推定結果がどう変化するか調べる「感度分析法」が採用されてきた (Rivest and Wells, 2001; De Uña-Álvarez and Veraverbeke, 2013, 2017; Emura and Hsu, 2020; Yeh et al., 2023; Emura et al., 2024)。例えば、コンピュータの型やパラメータ θ_i の選択に関

わらず検定が有意になる場合、この有意性の結論は様々な従属打ち切り構造下でも妥当といえる。一方、特定の θ_i の場合のみ有意になる結果などは、決定的なものにはならない。このような、様々な従属打ち切りのモデル下での感度分析は、従来の独立打ち切りの仮定下のみでの分析よりも有意性について厳しい条件を要求する。

感度分析を行う上で認識すべき点として、推定目標である治療効果 p は、打ち切りの分布や従属構造に依存せず生存関数のみで定まる母数である（式(2.1)および式(2.2)を参照）ことである。よって、様々なコピュラの仮定で変わる推定結果は、推定値（推定法）が変わる結果であり、母数が変わるわけではない。

4. ソフトウェア

3 節で紹介した手法は R パッケージ *compound.Cox* (Emura et al., 2019) 内の R 関数 *surv.factorial()* で実行可能である。この関数は、推定値 \hat{p}^{CG} とその SE、および p の 95% 信頼区間を算出する。また、検定統計量 F_N^{CG} とその棄却限界値 $c_{N,\alpha}$ 、および P 値も算出する。棄却限界はシミュレーション法と解析的手法の 2 つで計算され、それぞれ有意水準 $\alpha = 0.10, 0.05, 0.01$ の値を算出する。仮説 $H_0: C\mathbf{p} = \mathbf{0}$ における対比行列 C は自由に指定できるが、指定がなれば大域的仮説 $H_0: p_1 = p_2 = \dots = p_d$ を検定する。また、最大フォローアップ期間 τ も指定できるが、指定がなければ CG 推定量 $\hat{S}_i^{\text{CG}}(t)$ が全ての i と $t \in [0, \tau]$ で定義可能となるよう

$$\tau = \min_i \left[\max_j (X_{ij}) \right] < \min_i [\sup \{u: G_i(u) S_i(u) > 0\}]$$

をデフォルトにしている。これはいずれかの群で生存者が残っている最大の時間である。しかしながら、推定量 $\hat{S}_i^{\text{CG}}(t)$ が時間 τ の近傍で不安定になりやすいことから、 τ をさらに小さくとり、

$$(4.1) \quad \tau = \min_i \left[\max_{j: \delta_{ij}=1} (X_{ij}) \right] \leq \min_i \left[\max_j (X_{ij}) \right]$$

と指定でも良い（6 節の例を参照）。これはいずれかの群で死亡例が残っている最大の時間である。

入出力系統などの詳細については、オンラインマニュアルを参照頂きたい。利便性の都合上、*surv.factorial()* 関数では群間でコピュラが同一である条件下で計算する。即ち、 $C_{1\theta_1} = \dots = C_{d\theta_d}$ かつ $\theta_1 = \dots = \theta_d$ を仮定するため、コピュラの型 1 つとパラメータ 1 つを指定すれば十分である。

R コードの説明のために、R の *survival* パッケージ内にある *colon* データの解析コードの例を下に示す。まず入力すべき変数を準備し、“*surv.factorial(t.vec, d.vec, group, copula=CG, Clayton, theta)*” を実行すれば良い。ここでは “*copula=CG.Clayton*” でクレイトンコピュラを指定し “*theta*” でパラメータを指定している。3 群のデータの例なので、クレイトンコピュラのパラメータ $\theta_1 = \theta_2 = \theta_3 = 2$ の下で解析が行われる。局所的仮説の検定を行うには、対比行列 “*C=C12*” などと設定すれば良い。

```
library(survival)
library(compound.Cox)
data(cancer)
dat=subset(colon,etype==1)
## Treatment effects ##
```

```

> surv.factorial(t.vec,d.vec,group,copula=CG.Clayton,theta,t.upper=tau)
$copula.parameter
[1] 2

$p
  estimate      se    lower    upper      P
[1,] 0.4693207 0.01281679 0.4442003 0.4944412 9.916598e-01
[2,] 0.4701347 0.01283573 0.4449771 0.4952922 9.900103e-01
[3,] 0.5616860 0.01243152 0.5373207 0.5860513 3.487346e-07

$Var
      [,1]      [,2]      [,3]
[1,] 1.642701e-04 -8.720548e-05 -7.708513e-05
[2,] -8.720548e-05 1.647558e-04 -7.749042e-05
[3,] -7.708513e-05 -7.749042e-05 1.545426e-04

$F
[1] 11.65897

$c.simu
  0.10      0.05      0.01
2.418888 3.072035 4.776394

$c.anal
  0.10      0.05      0.01
2.303544 2.997522 4.609088

$p.value
  simu      anal
0.000000e+00 8.766784e-06

```

図 1. `surv.factorial()` 関数で `colon` データを解析した R コードのアウトプット.

```

t.vec=dat$time
d.vec=dat$status
trt=dat$rx
C12=matrix(c(1,-1,0), 1,3,byrow=TRUE)
C13=matrix(c(1,0,-1), 1,3,byrow=TRUE)
C23=matrix(c(0,1,-1), 1,3,byrow=TRUE)
group=as.numeric(trt) # 1=Obs; 2=Lev; 3=Lev+5FU
tau=min(tapply(t.vec[d.vec==1],group[d.vec==1], max)) ## maximum follow-up date
theta=2 ## copula parameter
surv.factorial(t.vec,d.vec,group,copula=CG.Clayton,theta,t.upper=tau)
surv.factorial(t.vec,d.vec,group,copula=CG.Clayton,theta,C=C12,t.upper=tau)
surv.factorial(t.vec,d.vec,group,copula=CG.Clayton,theta,C=C13,t.upper=tau)
surv.factorial(t.vec,d.vec,group,copula=CG.Clayton,theta,C=C23,t.upper=tau)

```

R コード出力は図 1 のようになる。推定値 $\hat{p}_i^{\text{CG}}, i = 1, 2, 3$ とそれらの SE, および p_i , $i = 1, 2, 3$ の 95% 信頼区間が表示されているのがわかる。加えて、ジャックナイフ法による $(\hat{p}_1^{\text{CG}}, \hat{p}_2^{\text{CG}}, \hat{p}_3^{\text{CG}})$ の共分散行列 $V(p^{\text{CG}})$ もある。また、統計量 F_N^{CG} と棄却限界値 $c_{N,\alpha}$ (シミュレーション法と解析的方法での 10%, 5%, 1% 棄却限界), および P 値が出力されているのがわかる。

5. シミュレーション

Emura et al. (2024) のシミュレーション研究により, `surv.factorial()` 関数を用いた推定・検定の性能はある程度調べられている。この既出のシミュレーションではまだ検証されていなかった局所的仮説の検定に関する性能, および大域的仮説検定と局所的仮説検定の関連を調べるために, 新たなシミュレーション実験を以下で行う。

5.1 シミュレーション設計

3つの治療群($d = 3$)を考え、各群 $i \in 1, 2, 3$ のサンプルサイズは $n_i = 50$ または $n_i = 100$ とした。生存時間と打ち切り時間のペア (T_{ij}, U_{ij}) , $i = 1, 2, 3$, $j = 1, 2, \dots, n_i$ をクレイトンコピュラモデルから生成し、コピュラパラメータを $\theta_i = 2$ ($\tau_{\theta_i} = 0.5$; 中程度の正相関)とした。周辺分布は T_{ij} が強度パラメータ λ_i の指数分布, U_{ij} が強度パラメータ μ_i の指数分布とし、フォローアップの最大時刻 τ に対して、 $U_{ij} = \min(U_{ij}, \tau)$ と再定義した。打ち切り時刻の強度パラメータを $\mu_1 = \mu_2 = \mu_3 = 1$, フォローアップ最大時刻を $\tau = 1$ とした。これより右打ち切り下での観測データ $\{(X_{ij}, \delta_{ij}); i = 1, 2, \dots, d, j = 1, 2, \dots, n_i\}$, $X_{ij} = \min(T_{ij}, U_{ij})$, $\delta_{ij} = 1_{T_{ij} \leq U_{ij}}$ を得た。

3.3 節で紹介した検定が有用となるのは、3つの群の差がKM推定量のグラフでは視覚的に判定しにくい状況である。シミュレーションではそのようなシナリオを2つ考えた：

- シナリオ 1 ($\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = 1.25$)：群1と群2に差が無いが、群1と群3および群2と群3に小さい差がある場合。このときの治療効果は、 $p_1 = 0.517$, $p_2 = 0.517$, $p_3 = 0.467$ 。
- シナリオ 2 ($\lambda_1 = 0.75$, $\lambda_2 = 1$, $\lambda_3 = 1.25$)：全ての群に小さい差がある場合。このときの治療効果は、 $p_1 = 0.556$, $p_2 = 0.497$, $p_3 = 0.447$ 。

値 1.25 や 0.75 は弱い治療効果の差を表す。例えば、強度 1.25 の群は強度 1.00 の群からハザードが 25% 増えるが、ハザード比 1.25 は非劣勢試験のマージンとして設定される典型的な量である (Yamada et al., 2018; Derkx et al., 2022)。

打ち切り割合 $P(U_{ij} < T_{ij})$ は1つ目のシナリオで、0.63(群1), 0.63(群2), 0.54(群3)であり、2つ目のシナリオで、0.73(群1), 0.63(群2), 0.54(群3)である。

以下の4つの検定問題を考える：

$$H_0: p_1 = p_2 \quad \text{vs.} \quad H_1: p_1 \neq p_2,$$

$$H_0: p_1 = p_3 \quad \text{vs.} \quad H_1: p_1 \neq p_3,$$

$$H_0: p_2 = p_3 \quad \text{vs.} \quad H_1: p_2 \neq p_3,$$

$$H_0: p_1 = p_2 = p_3 \quad \text{vs.} \quad H_1: p_1 \neq p_2, \quad \text{or} \quad p_1 \neq p_3, \quad \text{or} \quad p_2 \neq p_3.$$

最初の3つの局所的な検定では、対比行列を $\mathbf{C} = (1, -1, 0)$, $\mathbf{C} = (1, 0, -1)$, $\mathbf{C} = (0, 1, -1)$ とし、最後の大域的な検定では $\mathbf{C} = \mathbf{I}_3 - \mathbf{1}_3 \mathbf{1}'_3 / 3$ とした。各設定下で、データから統計量 F_N^{CG} と上側 5% 点 $c_{N,\alpha=0.05}$ を算出した。これを 1000 回繰り返し、 F_N^{CG} の平均値と検定の棄却率を計算した。

5.2 シミュレーション結果

表1にシミュレーションの結果を示す。帰無仮説が成立している $H_0: p_1 = p_2$ の検定では、 F_N^{CG} の平均値は 1 に近くなっ。これは漸近帰無分布の期待値が 1 であることと整合している。また、シナリオ 1 における $H_0: p_1 = p_2$ の検定では棄却率が名目値の 0.05 に近くなり、これは棄却限界値 $c_{N,\alpha}$ の設定法が妥当で、検定のサイズが正しく保たれていることを示している。帰無仮説が成立していない下では、 F_N^{CG} の平均値は 1 より大きくなり、近棄却率は 0.05 より高くなっ。特に、サンプルサイズが大きい($n_i = 100$)場合は、小さい場合($n_i = 50$)に比べ、棄却率は高くなっ。棄却限界値 $c_{N,\alpha}$ の 2 とおりの計算法(シミュレーション法と解析的方法)の間に差は殆ど見られなかっ。これは、解析的方法で用いた漸近帰無分布をカイ二乗分布で近似する手法が適切に働いていることを意味する。

シナリオ 1 では、大域的仮説($H_0: p_1 = p_2 = p_3$)は、局所的仮説($H_0: p_1 = p_3$ と $H_0: p_2 = p_3$)

表 1. シミュレーション 1000 回による統計量 F_N^{CG} の平均値および名目上の有意水準 $\alpha = 0.05$ の下での仮説 H_0 の棄却率.

| シナリオ 1* | $p_1 = p_2$ | n_i | $E[F_N^{\text{CG}}]$ | 棄却率: | 棄却率: | 棄却率の解釈 |
|----------|-------------------|-------|----------------------|----------|-------|--------|
| | | | | シミュレーション | 解析的 | |
| シナリオ 1* | $p_1 = p_2$ | 50 | 1.082 | 0.061 | 0.064 | 有意水準 |
| | | 100 | 0.962 | 0.044 | 0.047 | 有意水準 |
| シナリオ 2** | $p_1 = p_3$ | 50 | 1.749 | 0.130 | 0.128 | 検出力 |
| | | 100 | 2.032 | 0.172 | 0.175 | 検出力 |
| シナリオ 2** | $p_2 = p_3$ | 50 | 1.687 | 0.131 | 0.131 | 検出力 |
| | | 100 | 1.988 | 0.167 | 0.167 | 検出力 |
| シナリオ 2** | $p_1 = p_2 = p_3$ | 50 | 1.467 | 0.118 | 0.121 | 検出力 |
| | | 100 | 1.634 | 0.158 | 0.159 | 検出力 |
| シナリオ 2** | $p_1 = p_2$ | 50 | 1.855 | 0.146 | 0.144 | 検出力 |
| | | 100 | 2.553 | 0.233 | 0.230 | 検出力 |
| シナリオ 2** | $p_1 = p_3$ | 50 | 4.036 | 0.381 | 0.382 | 検出力 |
| | | 100 | 6.210 | 0.628 | 0.623 | 検出力 |
| シナリオ 2** | $p_2 = p_3$ | 50 | 1.680 | 0.133 | 0.131 | 検出力 |
| | | 100 | 1.992 | 0.165 | 0.167 | 検出力 |
| シナリオ 2** | $p_1 = p_2 = p_3$ | 50 | 2.447 | 0.312 | 0.308 | 検出力 |
| | | 100 | 3.540 | 0.515 | 0.514 | 検出力 |

*シナリオ 1: $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 1.25; p_1 = 0.517, p_2 = 0.517, p_3 = 0.467$;

**シナリオ 2: $\lambda_1 = 0.75, \lambda_2 = 1, \lambda_3 = 1.25; p_1 = 0.556, p_2 = 0.497, p_3 = 0.447$

よりも、わずかに棄却されにくくなった. これは、大域的仮説がその一部に検出不能な $p_1 = p_2$ を含むからである. 逆に、シナリオ 2 では、大域的仮説は、局所仮説 $H_0: p_1 = p_2$ と $H_0: p_2 = p_3$ よりも棄却されやすかったが、これは、 p_1, p_2, p_3 の 3 つとも異なる値を持つため、大域的仮説が棄却されやすいのは自然な結果である. 一方、シナリオ 2 における局所仮説 $H_0: p_1 = p_3$ は、大域的仮説よりもさらに棄却されやすかったが、これは p_1 と p_3 のペアは最も差が大きいペアだからである.

以上のシミュレーションの結果により、3.3 節で紹介した検定は、3 群の比較において妥当な有意水準(サイズ)と自然な検出力特性を持つことが示された. 検出力の性能では、群間差の大きいペアの局所的検定の検出力は、大域的検定の検出力より高まることが示された. しかしながら、通常の検定手順としては、最初に大域的検定を行って棄却された後に局所的検定を行うべきである.

6. データ解析

結腸がんのデータ解析例を通じて、3 節で紹介した手法を解説する. 解析は `surv.factorial()` 関数を用いて行う. 本節の解析結果を再現する R コードはオンライン Supplement を参照されたい.

Moertel et al. (1995) は、 $N = 929$ 人のステージ 3 結腸がんの切除手術を受けた被験者を次の 3 つの治療群($d = 3$)に無作為割付けし、各治療のがん再発に与える効果を検証した.

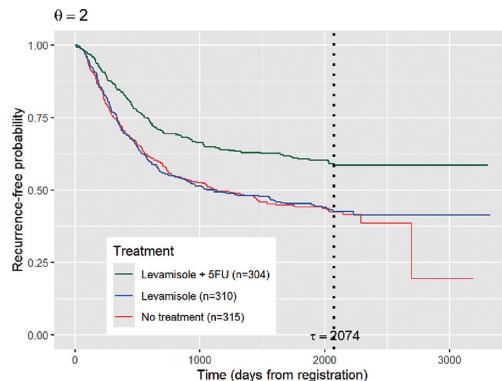


図 2. 結腸がんの切除手術を受けた 929 人の被験者の各治療における無再発確率の CG 推定値. クレイトンコピュラに $\theta_i = 2$ を適用. 経過観察(No adjuvant treatment), レバミゾール投与(Levamisole), レバミゾールと 5FU の併用(Levamisole + 5FU)の 3 群比較.

- 経過観察群 ($i = 1$) : 術後経過観察のみ,
- レバミゾール群 ($i = 2$) : レバミゾール投与による術後補助化学療法,
- 併用群 ($i = 3$) : レバミゾールと 5FU を併用した術後補助化学療法.

これらの治療が“無再発期間(Recurrence free interval)”に与える効果を推定・検定する. 無再発期間とは, 登録日から再発が確認されるまでの日数(生存時間 T_i に対応する数値)であるので, 長い期間ほど良い治療効果と見做す. 打ち切り時間は再発が観測されないまま観察が終わるまでの日数(打ち切り時間 U_i に対応する数値)であるが, 観察終了の理由は死亡によるものも含む(Moertel et al., 1995, Fig. 1). 再発時間と死亡時間は正に相関している可能性は十分あり, この独立打ち切りの仮定は妥当でない可能性がある(7 節で再び議論する). このような理由から, コピュラに基づく要因分析の手法(3 節)を用いて, 従属打ち切りを想定した解析を以下で行う. データは R パッケージ *survival* のデータセット *colon* から入手可能である.

図 2 に各群における無再発確率の CG 推定値(生存曲線)を図示した. ここで CG 推定値は, クレイトンコピュラにパラメータ $\theta_i = 2$ ($\tau_{\theta_i} = 0.5$; 中程度の正相関)を仮定した下で計算した. 図より, 併用群は他の 2 つの群に比べて無再発率が高い(治療効果が高い)ことが分かる. メディアン生存時間で見ると, レバミゾール投与群と経過観察群で 1200 日(3.3 年)ほどであるが, 併用群では 3000 日(8.2 年)以上もある. 一方, レバミゾール群と経過観察群での優劣は図から判定困難であり, 検定による判断に委ねることにする.

表 2 に群間差を検定した結果を示す. 3 つの群の生存率が比較可能な期間を $t \in [0, \tau]$ とし, ここに日数 $\tau = 2074$ は式(4.1)を用いた. 大域的な帰無仮説 $H_0: p_1 = p_2 = p_3$ の検定統計量は $F_N^{CG} = 11.659$ となり, これは帰無平均 1 から大きく離れた値であり, 有意水準 1% で有意となった($P < 0.001$). また, 併用群との比較に焦点を当てた局所的な帰無仮説($H_0: p_1 = p_3$ と $H_0: p_2 = p_3$)も 1% で有意となった($P < 0.001$). この有意差は, 図 2 から観測できる無再発率の大きな差から導かれる自然な結果である. 一方, 図 2 から予想できるよう, 経過観察群とレバミゾール群の同等性($H_0: p_1 = p_2$)は棄却されなかった. 以上の検定結果は, コピュラパラメータを $\theta_i = 2$ と仮定した下で行ったが, 他のコピュラパラメータ($\theta_i = 0.001, \theta_i = 4, \theta_i = 8$)の下での感度分析の結果も同じ結論となった(詳細は付録 3). 以下で個々の治療効果の推定を行う.

表 2. 結腸がんの切除手術を受けた 929 人の被験者の治療効果の同等性の検定結果. クレイトンコピュラにパラメータを $\theta_i = 2 (\tau_{\theta_i} = 0.5$; 中程度の正相関)とした下で計算.

| 比較する群 | 帰無仮説 H_0 | 検定統計量: | | P 値: | P 値: |
|-------------|-------------------|------------|-----------|--------|------|
| | | F_N^{CG} | シミュレーション法 | | |
| 経過観察とレバミゾール | $p_1 = p_2$ | 0.001 | 0.976 | 0.971 | |
| 経過観察と併用 | $p_1 = p_3$ | 18.037 | <0.001 | <0.001 | |
| レバミゾールと併用 | $p_2 = p_3$ | 17.672 | <0.001 | <0.001 | |
| 3 つの群全て | $p_1 = p_2 = p_3$ | 11.659 | <0.001 | <0.001 | |

表 3. 結腸がんの切除手術を受けた 929 人の被験者の治療効果の推定値. 母数 p_i は, i 番目の治療の相対的な治療効果を表し, 値が大きいほど高い治療効果を表す.

| コピュラパラメータ | 治療群 | 母数 | 推定値 | SE | 95%信頼区間 |
|--|--------|-------|-------|-------|--------------|
| $\theta_i = 0.001 (\tau_{\theta_i} = 0.0)$ | 経過観察 | p_1 | 0.469 | 0.013 | 0.444, 0.494 |
| | レバミゾール | p_2 | 0.470 | 0.013 | 0.446, 0.495 |
| | 併用 | p_3 | 0.561 | 0.012 | 0.538, 0.585 |
| $\theta_i = 2 (\tau_{\theta_i} = 0.5)$ | 経過観察 | p_1 | 0.469 | 0.013 | 0.444, 0.494 |
| | レバミゾール | p_2 | 0.470 | 0.013 | 0.445, 0.495 |
| | 併用 | p_3 | 0.562 | 0.012 | 0.537, 0.586 |
| $\theta_i = 4 (\tau_{\theta_i} = 0.67)$ | 経過観察 | p_1 | 0.469 | 0.013 | 0.444, 0.495 |
| | レバミゾール | p_2 | 0.470 | 0.013 | 0.445, 0.496 |
| | 併用 | p_3 | 0.562 | 0.013 | 0.537, 0.586 |
| $\theta_i = 8 (\tau_{\theta_i} = 0.80)$ | 経過観察 | p_1 | 0.469 | 0.013 | 0.443, 0.495 |
| | レバミゾール | p_2 | 0.471 | 0.013 | 0.446, 0.497 |
| | 併用 | p_3 | 0.561 | 0.013 | 0.536, 0.586 |

表 3 に様々な θ_i の値の下で推定した治療効果 \hat{p}_i^{CG} をまとめた. 例えば, $\theta_i = 2$ での結果を見ると, 併用群の治療効果は $\hat{p}_3 = 0.562$ (CI: 0.537, 0.586) であるから, 併用群の被験者集団が全被験者集団よりも再発を遅らせられる確率は 56.2% である. これより, 併用治療は良い治療と結論できる. これより併用群の無再発期間は他の 2 つの群のそれに比べて長いことが分かり, これは図 2 から得られる印象と整合する. 一方, レバミゾール群と経過観察群の治療効果の推定値を見ると, これらの群の被験者集団が全被験者集団より無再発期間を長く出来る確率は 47% ほどであり, これらは良くない治療といえる. 以上のような結果と解釈は, θ_i を他の値に変えて解析してもほぼ同じであり(表 3), さらにゲンベルコピュラの下でも同じであった(結果は割愛した). 故に, 併用群が高い治療効果, レバミゾール群と経過観察群が低い治療効果を示す, という結論は従属打ち切りの構造に影響を受けない頑健なものであることが示された. 以上の結果と解釈は, 独立打ち切りの仮定下で得られた Moertel et al. (1995) の報告結果をより強固にするものである.

7. 結論と考察

本論文では、打ち切りが独立でない状況における生存時間の群間比較を行うための要因計画・要因分析法 (Emura et al., 2024) を詳説した。この方法はコピュラを用いて、打ち切り時間と生存時間の従属構造を特定した下で、治療効果の推定・検定を実行する。コピュラの型や相関パラメータの特定は、実際のデータからは困難であるため、コピュラとパラメータの候補をいくつか用意し、それらの下での解析結果を吟味する感度分析法を紹介した(3.3 節, 6 節)。また、本論文では新たなシミュレーション実験を行い、原著では考察されていなかった局所的検定の検出力特性および大域的検定との関連を明らかにした。最後に紹介した結腸がんデータの解析例も、原著論文には無い新たな例であり、提案手法の有用性を再確認した。スペースの制限により、本論文の数値例ではクレイトンコピュラの結果のみを報告したが、グンベルコピュラやフランクコピュラなどの他のアルキメデスコピュラでも同様の解析を行うことが可能となる。本稿ではクレイトンコピュラ下での結果のみ報告したが、グンベルやフランクコピュラ下で解析する場合は、Supplement の R コードで `copula=CG.Clayton` を `copula=CG.Gumbel` に変更などすれば良い。ただし、紹介した手法はアルキメデスコピュラによるモデルに限定されるため、アルキメデスコピュラでないコピュラ(例えば、正規コピュラ、FGM コピュラ、プラケットコピュラなどがある)をどう活用するのか考察することは今後の課題となる。

結腸がんのデータ解析では、再発時間と死亡時間は正に相関しているという考え方の下、独立打ち切りの仮定が妥当でない可能性を考慮した解析を行った。死亡と増悪(再発・転移など)が相関しているという考えは、がん領域では一般的であり、がん臨床データの解析では独立打ち切りの仮定を必要としない競合リスクモデルによる解析が推奨されることもある(齋藤・室谷, 2024)。競合リスク解析では、“原因別ハザードモデル”または“部分分布ハザードモデル”などを用いれば、リスク間の独立性が成立するかどうかに関わらず解析を進めることができある。一方、本論文の従属打ち切りのコピュラモデルは、競合リスクモデルの一種である“潜在故障時間モデル”(宮村, 1982; 武富・山本, 2023; 道前, 2023, 2025)と位置付けられる。潜在故障時間モデルにおいてコピュラを用いる場合、コピュラのパラメータをデータから推定する方法 (Emura and Chen, 2016; Deresa and Van Keilegom, 2020, 2022; Czado and Van Keilegom, 2023; Lo et al., 2024; Michimae et al., 2024; Delhelle and Van Keilegom, 2025) と、そうでない仮定されたコピュラの方法 (Zheng and Klein, 1994, 1995; De Uña-Álvarez and Veraverbeke, 2013, 2017; Emura and Michimae, 2017; Emura and Hsu 2020; Rivest and Wells, 2001; Michimae and Emura, 2022; Emura et al., 2024) がある。本論文の要因実験・要因分析の方法論では、Rivest and Wells (2001) に従い、仮定されたコピュラの下で CG 推定法を適用していることから、感度分析法(3.3 節, 6 節)を採用した。コピュラモデルは、がん臨床試験において無増悪生存期間と全生存期間の相関の推定に用いられることも多く(江村・大庭, 2024; Simon et al., 2025), 本論文のデータ解析で考えた再発時間と打ち切り時間の相関のコピュラモデルもそれに準じたものと理解できる。

謝 辞

改稿に当たり、有益なコメントを下さった 2 名の査読者と編集委員の方に感謝申し上げます。Emura et al. (2024) に共著者として重要な貢献を頂いた Marc Ditzhaus 氏は、本論文を執筆中の 2024 月 9 月に永眠されました。謹んで感謝と哀悼の意を表し、ご冥福をお祈りします。本研究は科研費(22K11948; 20H04147; 24K13473)の助成を受けました。

付録 1. クロネッカーリー積と対比行列による帰無仮説

任意の $m \times n$ 行列 \mathbf{D} と、任意の行列 \mathbf{E} に対して、クロネッカーリー積は次のように定義される：

$$\mathbf{D} \otimes \mathbf{E} = \begin{pmatrix} d_{11} & \dots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{m1} & \dots & d_{mn} \end{pmatrix} \otimes \mathbf{E} = \begin{pmatrix} d_{11}\mathbf{E} & \dots & d_{1n}\mathbf{E} \\ \vdots & \ddots & \vdots \\ d_{m1}\mathbf{E} & \dots & d_{mn}\mathbf{E} \end{pmatrix}.$$

これより、

$$\mathbf{A} = \mathbf{I}_d \otimes \left(\frac{\mathbf{1}'_d}{d} \right) = \frac{1}{d} \begin{pmatrix} 1, \dots, 1 & 0, \dots, 0 & \dots & 0, \dots, 0 \\ 0, \dots, 0 & 1, \dots, 1 & \dots & 0, \dots, 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0, \dots, 0 & 0, \dots, 0 & \dots & 1, \dots, 1 \end{pmatrix}$$

であるから、

$$\begin{aligned} \mathbf{A}\mathbf{w} &= \frac{1}{d} \begin{pmatrix} 1, \dots, 1 & 0, \dots, 0 & \dots & 0, \dots, 0 \\ 0, \dots, 0 & 1, \dots, 1 & \dots & 0, \dots, 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0, \dots, 0 & 0, \dots, 0 & \dots & 1, \dots, 1 \end{pmatrix} \\ &\quad \times (w_{11}, \dots, w_{1d} : w_{21}, \dots, w_{2d} : \dots : w_{d1}, \dots, w_{dd})' \\ &= \frac{1}{d} \begin{pmatrix} w_{11} + \dots + w_{1d} \\ \vdots \\ w_{d1} + \dots + w_{dd} \end{pmatrix} = \begin{pmatrix} p_1 \\ \vdots \\ p_d \end{pmatrix} \end{aligned}$$

となり、 $\mathbf{p} = \mathbf{A}\mathbf{w}$ が示された。ちにみに、クロネッカーリー積は可換でなく、 $\mathbf{D} \otimes \mathbf{E} = \mathbf{E} \otimes \mathbf{D}$ は一般に成立しないため、積の順序に注意する必要がある。

また、対比行列を $\mathbf{C} = \mathbf{P}_d = \mathbf{I}_d - \mathbf{1}_d\mathbf{1}'_d/d$ とおくと、帰無仮説 $H_0: \mathbf{C}\mathbf{p} = \mathbf{0}$ は次のようになる：

$$\mathbf{C}\mathbf{p} = \begin{pmatrix} 1 - 1/d & -1/d & \dots & -1/d \\ -1/d & 1 - 1/d & \dots & -1/d \\ \vdots & \ddots & \ddots & \vdots \\ -1/d & -1/d & \dots & 1 - 1/d \end{pmatrix} \times \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_d \end{pmatrix} = \begin{pmatrix} p_1 - \bar{p} \\ p_2 - \bar{p} \\ \vdots \\ p_d - \bar{p} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

これより、帰無仮説は $H_0: p_1 = p_2 = \dots = p_d$ と書き換えられる。

付録 2. 漸近論の正則条件

Emura et al. (2024) の付録 A.3 にある 3 つの条件を要する。1 つ目は、 $N \rightarrow \infty$ のとき、ある定数 $\lambda_i \in (0, 1)$ が存在し、

$$\frac{\mathbf{n}'}{N} = \frac{(n_1, n_2, \dots, n_d)}{N} \longrightarrow \boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_d), \quad i = 1, 2, \dots, d,$$

となることであり、2 つ目は $\pi(\tau) = P(X_{ij} \geq \tau) > 0$ となる $\tau > 0$ が存在すること、3 つ目は関数 $\phi'_{i\theta_i}(t) = d\phi_{i\theta_i}(t)/dt$ と $\psi'_{i\theta_i}(t) = d\psi_{i\theta_i}(t)/dt$ が $(\pi(\tau), 1)$ 上で有界であることであり、ここに

$\psi_{i\theta_i}(t) = -u\phi'_{i\theta_i}(t)$ である。これら3つは緩やかな条件であり、特に3つ目の条件は、クレイトン・フランク・グンベルコピュラの全てで満足している。

付録3. 結腸がんのデータ解析の感度分析

結腸がんのデータ解析例における感度分析の結果を示す。すなわち、クレイトンコピュラのパラメータを $\theta_i = 0.001, \theta_i = 2, \theta_i = 4, \theta_i = 8$ と変化させたときの解析結果を調べる。図3に従属の強さを変えたときの生存関数のCG推定値を与えた。この結果によると、 θ_i の値が大きくなるほど、再発率の推定値は、どの群でも少しずつ減少していくことが分かる。特に、クレイトンコピュラの下側裾従属性からか、2000日以降の生存時間に大きな群間差がみられる。よって、再発率の推定値は、打ち切りの従属の影響を少なからず受けていることが分かる。一方で、群間比較の検定・推定の結果は θ_i の値の影響を殆ど受けないことが分かる(表4)。この結果より、従属打ち切りは、再発率自体には影響しても、再発率の群間差にはあまり影響を受けないと結論できる。

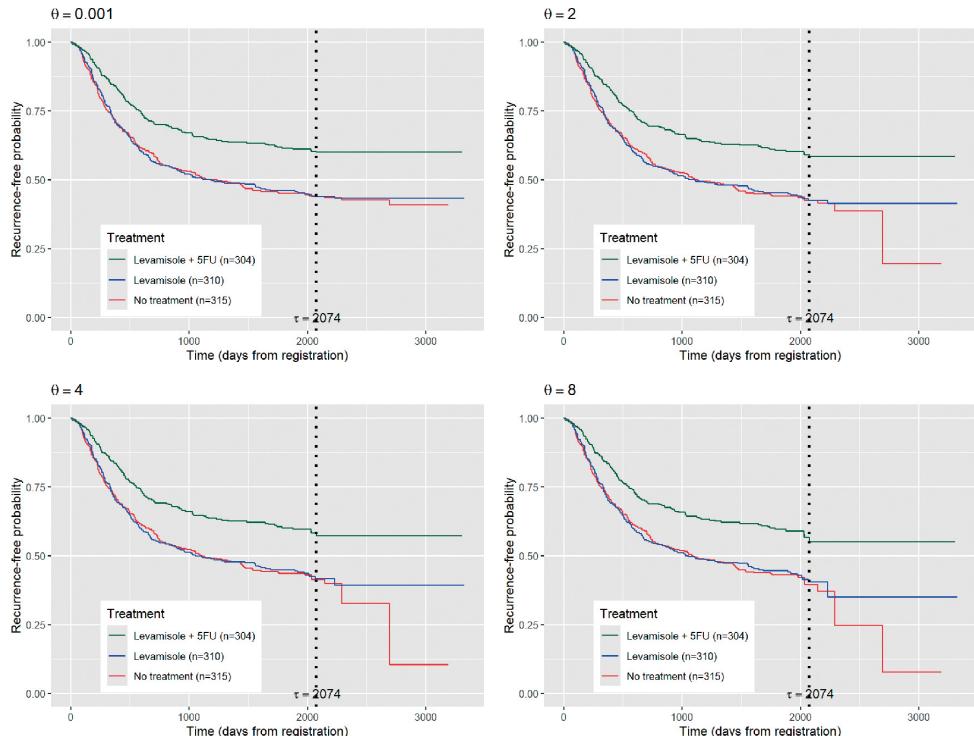


図3. 結腸がんの切除手術を受けた929人の被験者の各治療における無再発確率のCG推定値。クレイトンコピュラのパラメータ θ_i に4つの値を適用。術後経過観察(No adjuvant treatment), レバミゾール投与(Levamisole), レバミゾールと5FUの併用(Levamisole + 5FU)の比較。

表 4. 結腸がんの切除手術を受けた 929 人の被験者の治療効果の差の検定結果. クレイトン
コピュラのパラメータ θ_i に 4 つの値を与え感度分析を行った.

| θ_i | 比較する群 | 帰無仮説 H_0 | 統計量: | P 値: | P 値: |
|------------|-------------|-------------------|-------------------|----------|--------|
| | | | F_N^{CG} | シミュレーション | 解析的 |
| 0.001 | 経過観察とレバミゾール | $p_1 = p_2$ | 0.003 | 0.959 | 0.958 |
| | 経過観察と併用 | $p_1 = p_3$ | 18.566 | <0.001 | <0.001 |
| | レバミゾールと併用 | $p_2 = p_3$ | 18.069 | <0.001 | <0.001 |
| | 3 つ全て | $p_1 = p_2 = p_3$ | 11.921 | <0.001 | <0.001 |
| 2 | 経過観察とレバミゾール | $p_1 = p_2$ | 0.001 | 0.976 | 0.971 |
| | 経過観察と併用 | $p_1 = p_3$ | 18.037 | <0.001 | <0.001 |
| | レバミゾールと併用 | $p_2 = p_3$ | 17.672 | <0.001 | <0.001 |
| | 3 つ全て | $p_1 = p_2 = p_3$ | 11.659 | <0.001 | <0.001 |
| 4 | 経過観察とレバミゾール | $p_1 = p_2$ | 0.003 | 0.951 | 0.957 |
| | 経過観察と併用 | $p_1 = p_3$ | 17.616 | <0.001 | <0.001 |
| | レバミゾールと併用 | $p_2 = p_3$ | 17.139 | <0.001 | <0.001 |
| | 3 つ全て | $p_1 = p_2 = p_3$ | 11.381 | <0.001 | <0.001 |
| 8 | 経過観察とレバミゾール | $p_1 = p_2$ | 0.011 | 0.905 | 0.915 |
| | 経過観察と併用 | $p_1 = p_3$ | 16.715 | <0.001 | <0.001 |
| | レバミゾールと併用 | $p_2 = p_3$ | 15.939 | <0.001 | <0.001 |
| | 3 つ全て | $p_1 = p_2 = p_3$ | 10.759 | <0.001 | <0.001 |

参考文献

Andersen, P. K. and Perme, M. P. (2010). Pseudo-observations in survival analysis, *Statistical Methods in Medical Research*, **19**(1), 71–99.

Brunner, E. and Puri, M. L. (2001). Nonparametric methods in factorial designs, *Statistical Papers*, **42**(1), 1–52.

Brunner, E., Konietzschke, F., Pauly, M. and Puri, M. L. (2017). Rank-based procedures in factorial designs: Hypotheses about non-parametric treatment effects, *Journal of the Royal Statistical Society: Series B*, **79**(5), 1463–1485.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, **65**(1), 141–151.

Collett, D. (2023). *Modelling Survival Data in Medical Research*, 4th ed., Chapman & Hall, Oxfordshire, UK.

Czado, C. and Van Keilegom, I. (2023). Dependent censoring based on parametric copulas, *Biometrika*, **110**(3), 721–738.

Das, S., Bhattacharya, R. and Shome, M. (2024). A multi-treatment two stage allocation design for survival outcomes with nonrandom censoring: A copula-based approach, *Sequential Analysis*, **43**(4), 461–476.

De Uña-Álvarez, J. and Veraverbeke, N. (2013). Generalized copula-graphic estimator, *Test*, **22**(2), 343–360.

De Uña-Álvarez, J. and Veraverbeke, N. (2017). Copula-graphic estimation with left-truncated and right-censored data, *Statistics*, **51**(2), 387–403.

Delhelle, M. and Van Keilegom, I. (2025). Copula based dependent censoring in cure models, *Test*, <https://doi.org/10.1007/s11749-024-00961-7>.

Deresa, N. W. and Van Keilegom, I. (2020). A multivariate normal regression model for survival data subject to different types of dependent censoring, *Computational Statistics & Data Analysis*, **144**, 106879.

Deresa, N. W. and Van Keilegom, I. (2025). Semiparametric transformation models for survival data with dependent censoring, *Annals of the Institute of Statistical Mathematics*, <https://doi.org/10.1007/s10463-024-00921-w>.

Deresa, N. W., Van Keilegom, I. and Antonio, K. (2022). Copula-based inference for bivariate survival data with left truncation and dependent censoring, *Insurance: Mathematics and Economics*, **107**, 1–21.

Derkzen, J. W., Smit, K. C., May, A. M. and Punt, C. J. (2022). Systematic review and non-inferiority meta-analysis of randomised phase II/III trials on S-1-based therapy versus 5-fluorouracil-or capecitabine-based therapy in the treatment of patients with metastatic colorectal cancer, *European Journal of Cancer*, **166**, 73–86.

Dettoni, R., Marra, G. and Radice, R. (2020). Generalized link-based additive survival models with informative censoring, *Journal of Computational and Graphical Statistics*, **29**(3), 503–512.

Ditzhaus, M., Dobler, D. and Pauly, M. (2021). Inferring median survival differences in general factorial designs via permutation tests, *Statistical Methods in Medical Research*, **30**(3), 875–891.

Ditzhaus, M., Genuneit, J., Janssen, A. and Pauly, M. (2023). CASANOVA: Permutation inference in factorial survival designs, *Biometrics*, **79**(3), 203–215.

Dobler, D. and Pauly, M. (2018). Bootstrap- and permutation-based inference for the Mann-Whitney effect for right-censored and tied data, *Test*, **27**(3), 639–658.

Dobler, D. and Pauly, M. (2020). Factorial analyses of treatment effects under independent right-censoring, *Statistical Methods in Medical Research*, **29**(2), 325–343.

Dobler, D., Friedrich, S. and Pauly, M. (2020). Nonparametric MANOVA in meaningful effects, *Annals of the Institute of Statistical Mathematics*, **72**(4), 997–1022.

Durante, F. and Sempi, C. (2016). *Principles of Copula Theory*, CRC Press, Boca Raton, Florida.

Efron, B. (1967). The two sample problem with censored data, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol.5.4, 831–853.

江村剛志 (2023). 生存時間の2標本問題—コピュラに基づく従属打ち切り問題への対処—, 日本統計学会誌, **52**(2), 295–317.

江村剛志 (2025). 『コピュラ理論の基礎』, コロナ社, 東京.

Emura, T. and Chen, Y. H. (2016). Gene selection for survival data under dependent censoring, a copula-based approach, *Statistical Methods in Medical Research*, **25**(6), 2840–2857.

Emura, T. and Chen, Y. H. (2018). *Analysis of Survival Data with Dependent Censoring: Copula-Based Approaches*, Springer, Singapore.

Emura, T. and Hsu, J. H. (2020). Estimation of the Mann-Whitney effect in the two-sample problem under dependent censoring, *Computational Statistics & Data Analysis*, **150**, 106990.

Emura, T. and Michimae, H. (2017). A copula-based inference to piecewise exponential models under dependent censoring, with application to time to metamorphosis of salamander larvae, *Environmental and Ecological Statistics*, **24**(1), 151–173.

江村剛志, 大庭幸治 (2024). 生存時間変数に対する代替性評価—メタアナリシスアプローチ—, 計量生物学, **45**(1), 67–85.

Emura, T., Matsui, S. and Chen, H. Y. (2019). compound. Cox: Univariate feature selection and compound covariate for predicting survival, *Computer Methods and Programs in Biomedicine*, **168**, 21–37.

Emura, T., Ditzhaus, M., Dobler, D. and Murotani, K. (2024). Factorial survival analysis for treatment effects under dependent censoring, *Statistical Methods in Medical Research*, **33**(1), 61–79.

Fizazi, K., Foulon, S., et al. (2022). Abiraterone plus prednisone added to androgen deprivation therapy and docetaxel in de novo metastatic castration-sensitive prostate cancer (PEACE-1): A multicentre, open-label, randomised, phase 3 study with a 2×2 factorial design, *The Lancet*, **399**(10336), 1695–1707.

Freidlin, B. and Korn, E. (2017). Two-by-Two factorial cancer treatment trials: is sufficient attention being paid to possible interactions?, *Journal of the National Cancer Institute*, **109**(9), djx146.

Futschik, A., Taus, T. and Zehetmayer, S. (2019). An omnibus test for the global null hypothesis, *Statistical Methods in Medical Research*, **28**(8), 2292–2304.

Gares, V., Hudson, M., Manuguerra, M. and Gebski, V. (2024). Effect of a correlated competing risk on marginal survival estimation in an accelerated failure time model, *Communications in Statistics—Simulation and Computation*, <https://doi.org/10.1080/03610918.2024.2380005>.

Gorfine, M., Schlesinger, M. and Hsu, L. (2020). K -sample omnibus non-proportional hazards tests based on right-censored data, *Statistical Methods in Medical Research*, **29**, 2830–2850.

Gottselig, M., Schneider, S., Demarqui, F. N. (2025). Cure rate regression models for dependent censoring under a copula-based approach, *Japanese Journal of Statistics and Data Science*, <https://doi.org/10.1007/s42081-025-00294-1>.

Jo, J. H., Gao, Z., Jung, I., et al. (2023). Copula graphic estimation of the survival function with dependent censoring and its application to analysis of pancreatic cancer clinical trial, *Statistical Methods in Medical Research*, **32**(5), 944–962.

Klein, J. P. and Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*, Springer, New York.

Koziol, J. A. and Jia, Z. (2009). The concordance index C and the Mann-Whitney parameter $\Pr(X > Y)$ with randomly censored data, *Biometrical Journal*, **51**(3), 467–474.

Li, D., Hu, X. J. and Wang, R. (2023). Evaluating association between two event times with observations subject to informative censoring, *Journal of the American Statistical Association*, **118**(542), 1282–1294.

Lo, S. M., Wilke, R. A. and Emura, T. (2024). A semiparametric model for the cause-specific hazard under risk proportionality, *Computational Statistics & Data Analysis*, **195**, 107953.

Lu, Z. and Zhang, W. (2012). Semiparametric likelihood estimation in survival models with informative censoring, *Journal of Multivariate Analysis*, **106**, 187–211.

松井茂之 (2023). 要因デザイン, 『臨床試験の事典』(丹後俊郎, 松井茂之 編), 5章 22節, 朝倉書店, 東京.

道前洋史 (2023). 競合リスクを伴う左側切断・右側打ち切りデータの解析—現状と今後の展望について, 日本統計学会誌, **52**(2), 203–220.

道前洋史 (2025). 相関した競合リスクを伴う左側切断・右側打ち切りデータのペイズ流解析, 統計数理, **73**(2), 167–188.

Michimae, H. and Emura, T. (2022). Likelihood inference for copula models based on left-truncated and competing risks data from field studies, *Mathematics*, **10**(13), 2163.

Michimae, H., Emura, T., Miyamoto, A. and Kishi, K. (2024). Bayesian parametric estimation based on left-truncated competing risks data under bivariate Clayton copula models, *Journal of Applied Statistics*, **51**(13), 2690–2708.

Mishra, P., Singh, U., Pandey, C. M., Mishra, P. and Pandey, G. (2019). Application of student's t-test, analysis of variance and covariance, *Annals of Cardiac Anaesthesia*, **22**(4), 407–411.

宮村鍛夫 (1982). 競合モデルに関する統計的手法, オペレーションズ・リサーチ, **27**, 871–876.

Moertel, C. G., Fleming, T. R., Macdonald, J. S., et al. (1995). Fluorouracil plus levamisole as effective adjuvant therapy after resection of stage III colon carcinoma: A final report, *Annals of Internal*

Medicine, **122**(5), 321–326.

Montgomery, D. C. (2019). *Design and Analysis of Experiments*, John Wiley & Sons, New Jersey.

Moradian, H., Larocque, D. and Bellavance, F. (2019). Survival forests for data with dependent censoring, *Statistical Methods in Medical Research*, **28**(2), 445–461.

Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd ed., Springer, New York.

Nowak, C. P., Mtze, T. and Konietschke, F. (2022). Group sequential methods for the Mann-Whitney parameter, *Statistical Methods in Medical Research*, **31**(10), 2004–2020.

Perme, M. P. and Manevski, D. (2019). Confidence intervals for the Mann-Whitney test, *Statistical Methods in Medical Research*, **28**(12), 3755–3768.

Rivest, L. P. and Wells, M. T. (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring, *Journal of Multivariate Analysis*, **79**(1), 138–155.

齋藤哲雄, 室谷健太 (2024). がん臨床試験と競合リスク・マルチステートモデル, 計量生物学, **45**(1), 37–65.

Schneider, S., Demarqui, F. N., Colosimo, E. A. and Mayrink, V. D. (2020). An approach to model clustered survival data with dependent censoring, *Biometrical Journal*, **62**(1), 157–174.

Schneider, S., Demarqui, F. and de Freitas Costa, E. (2022). Free-ranging dogs' lifetime estimated by an approach for long-term survival data with dependent censoring, *Environmental and Ecological Statistics*, **29**(4), 869–911.

Simon, F., Ligtvoet, R., Robrecht, S., et al. (2025). Endpoint surrogacy in first-line chronic lymphocytic leukemia, *Journal of Clinical Oncology*, **43**(4), 381–391.

Staplin, N., Kimber, A., Collett, D. and Roderick, P. (2015). Dependent censoring in piecewise exponential survival models, *Statistical Methods in Medical Research*, **24**(3), 325–341.

杉本知之, 田中健太 (2023). 2変量生存時間モデルにおけるコピュラとその利用, 日本統計学会誌, **52**(2), 153–176.

武富奈菜美, 山本和嬉 (2023). 生存時間解析・信頼性解析のための統計モデル, 日本統計学会誌, **52**(2), 69–112.

Tewari, K. S., Sill, M. W., et al. (2014). Improved survival with bevacizumab in advanced cervical cancer, *New England Journal of Medicine*, **370**(8), 734–743.

東京大学教養学部統計学教室 (1992). 『自然科学の統計学』, 第3章, 東京大学出版会, 東京.

Tsiatis, A. A. (1975). Nonidentifiability aspect of the problem of competing risks, *Proceedings of the National Academy of Sciences*, **72**, 20–22.

Xu, J., Ma, J., Connors, M. H. and Brodaty, H. (2018). Proportional hazard model estimation under dependent censoring using copulas and penalized likelihood, *Statistics in Medicine*, **37**(14), 2238–2251.

Yamada, Y., Denda, T., Gamoh, M., et al. (2018). S-1 and irinotecan plus bevacizumab versus mFOLFOX6 or CapeOX plus bevacizumab as first-line treatment in patients with metastatic colorectal cancer (TRICOLORE): A randomized, open-label, phase III, noninferiority trial, *Annals of Oncology*, **29**(3), 624–631.

Yeh, C. T., Liao, G. Y. and Emura, T. (2023). Sensitivity analysis for survival prognostic prediction with gene selection: A copula method for dependent censoring, *Biomedicines*, **11**(3), 797.

Zheng, M. and Klein, J. P. (1994). A self-consistent estimator of marginal survival functions based on dependent competing risk data and an assumed copula, *Communications in Statistics—Theory and Methods*, **23**(8), 2299–2311.

Zheng, M. and Klein, J. P. (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula, *Biometrika*, **82**(1), 127–138.

A Review of Factorial Survival Analysis for Treatment Effects
under Dependent Censoring
—A Copula-based Approach—

Takeshi Emura¹ and Kenta Murotani²

¹School of Informatics and Data Science, Hiroshima University

²School of Medical Technology, Kurume University

For estimating treatment effects on survival, it is not suitable to apply the classical analysis of variance method under the normal distribution model. Nonparametric methods are desirable for defining and estimating treatment effects, which do not impose any parametric model. In this article, we review a nonparametric method for factorial survival analysis under dependent censoring, which utilizes copula-graphic estimators. To implement the method, we introduce an R function, `surv.factorial()`. To investigate the performance of the global and local tests for no treatment effects, we conduct simulation studies and report the results. We also analyze a dataset on colon cancer patients and discuss the consequences.