Cox 比例ハザードモデルでの対数ハザード比と連続変数の関係を制限付き 3 次スプライン曲線で可視化する方法

SAS Institute Japan 株式会社 JMP ジャパン事業部

2023 年 8 月

このドキュメントは日本のテクニカルサポートに何度か問い合わせがあったことを受けて作成しています。問い合わせでは主にハザード比と連続変数の関係を制限付き 3 次スプライン曲線(自然 3 次スプライン曲線)で視覚化する方法について興味を持たれていました。また、ベースラインハザード関数$h_0(t)$に対するハザード比ではなく、連続変数が特定の値をとるときのハザード関数に対するハザード比を描きたいというご質問をいただくこともあります。

＜ご質問の例＞

- 連続変数とハザード比に関する制限付き 3 次スプライン(restricted cubic spline)のグラフを描きたいのですがどのようにしたらよいでしょうか。
- ハザード比を縦軸、調べたい因子（連続変数）を横軸にとって制限付き 3 次スプライン曲線を書きたいのですが JMP で解析することは可能でしょうか？

このドキュメントでは Heinzl and Kaider (1997)にて紹介されている方法のうち、連続な共変量を制限付き 3 次スプライン曲線の効果としたときの Cox 比例ハザードモデルをあてはめます。この時 JMP Pro の機能である[一般化回帰]を用いて、Cox 比例ハザードモデルをあてはめ、そこから得られる対数ハザードを用いてグラフ化する方法を紹介します。より具体的には Heinzl and Kaider (1997)の Fig.1 と同様のグラフ(縦軸は対数ハザード比、横軸は連続な共変量)を描くための方法です。なお、制限付き 3 次スプラインは自然 3 次スプラインと呼ばれることもあり、JMP では節点スプラインと呼んでいます。

1. Cox 比例ハザードモデルでの制限付き 3 次スプライン曲線の利用について

制限付き 3 次スプライン曲線を用いて、Cox 比例ハザードの前提を確認する方法が多くの研究で論じられています。Heinzl and Kaider (1997)では(1)連続な共変量の非線形性の検証、(2)時間と共変量の交互作用の検証、(3)時間と 2 水準の時間依存性共変量の交互作用の検証、という 3 つについて紹介されています。

まず、ハザードを$h(t;Z)$とすると、Cox 比例ハザードモデルは次の式で表せます。

$$h(t;Z) = h_0(t) \cdot \exp\big(f(Z;\beta)\big)$$

このとき(1)、(2)、(3)のそれぞれで想定されているモデルは下記のようなものです。

(1) $f(Z;\beta)$の部分をスプライン関数とするモデル

(2) $h_0(t)$の部分をスプライン関数とするモデル*

*スプライン関数を$g(t)$とすると、$h(t;Z) = g(t) \cdot \exp(Z\beta)$とするモデル

(3) Zも時間依存性共変量とするモデル

JMP では時間依存性共変量を扱うことができないため、**(2)、(3)のモデルはあてはめられません**。

以降 2.では**(1)のモデルをあてはめる方法**を紹介します。

2. JMP Pro の[一般化回帰]を用いて Cox 比例ハザードモデルをあてはめる**

**JMP Pro 17.1 を使用しています。バージョンによっては若干操作が異なることがあります。

ここでは JMP のサンプルデータ VA Lung Cancer.jmp を用います。データテーブルの「生存日数」を応答、「年齢」列を共変量として Cox 比例ハザードモデルをあてはめてみます。

＜操作:モデルのあてはめ＞

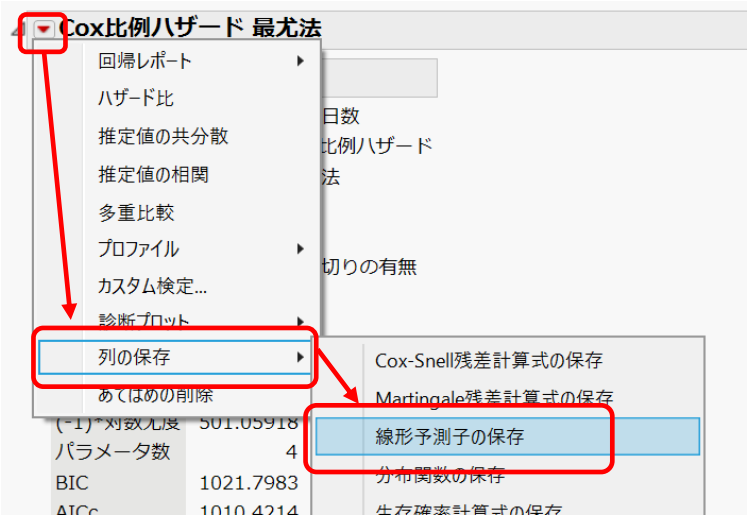
1. JMP の[ヘルプ]メニューより[サンプルの索引]を選び、表示されウィンドウで[サンプルデータフォルダを開く]よりデータテーブルを開きます。
2. データテーブルで[分析]>[モデルのあてはめ]を選択します。
3. モデルダイアログで次のように設定します。

   <u>手法</u>:一般化回帰、<u>分布</u>:Cox 比例ハザード、<u>打ち切り</u>:打ち切りの有無、

   <u>打ち切りの値</u>:Censored、<u>Y</u>:生存日数、<u>モデルの効果</u>:年齢
4. モデルの効果に追加した「年齢」を選び、[属性]>[節点スプライン効果]を選びます。
5. 節点の数を入力するウィンドウではそのまま[OK]をクリックします(デフォルトの設定を使用します)。

   これにより効果が"**年齢&節点スプライン**"となり、制限付き 3 次スプライン曲線の効果が指定されます。

6. [実行]をクリックします。
7. レポートウィンドウが表示されますので、「Cox 比例ハザード 最尤法」と書かれた隣の赤い三角形をクリックして[列の保存]>[線形予測子の保存]を選びます。



以上の操作でデータテーブルに「線形予測子」列が保存されますが、この列はスプライン関数とした$f(Z;\beta)$が計算式として設定されています。$f(Z;\beta)$は基準を$h_0(t)$とした対数ハザード比ですので、その指数をとることで、ハザード比となります。

＜(オプショナル)操作：対数ハザード比のハザード比への変換＞

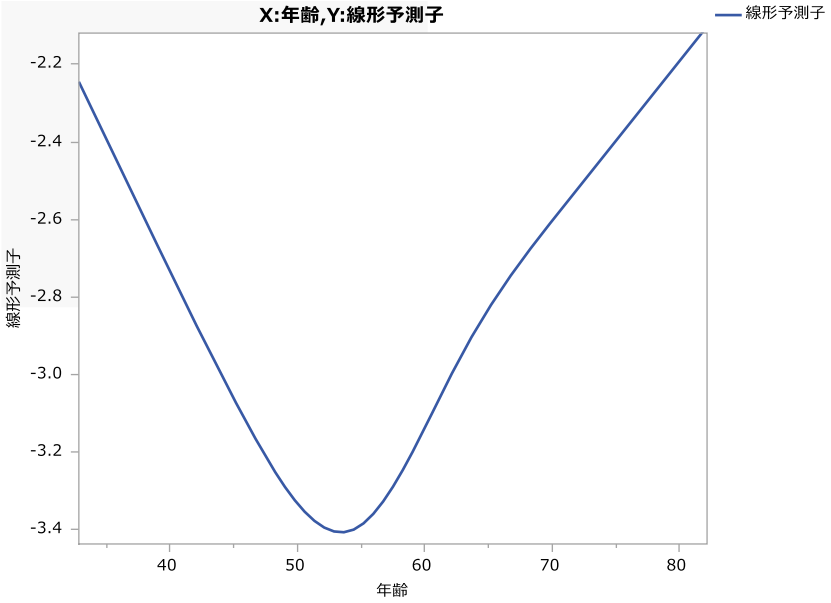
1. データテーブルで「線形予測子」列を選択し、列名上で右クリックから[計算式列の新規作成]>[対数]>[指数]を選びます。
2. この操作により「指数[線形予測子]」列が保存されることを確認します。

3. 対数ハザード比を視覚化する

ここでは JMP の[グラフビルダー]を用いて、縦軸が対数ハザード比、横軸が年齢のグラフを作成します。

＜操作：[グラフビルダー]による可視化＞

1. [グラフ]>[グラフビルダー]を選択します。
2. [Y]に「線形予測子」列、[X]に「年齢」列を指定します。
3. 上部アイコンにおいて、右から 3 番目の「f(.)」をクリックします。
4. [終了]をクリックします。



注）JMPではこの曲線の信頼区間を描画することはできません。

4. 基準(リファレンス)を変更するとき

2、3で描画したグラフは基準を$h_0(t)$とした対数ハザード比です。場合によっては共変量を任意の値としたときのハザードを基準とすることもあるかと思います。例えば2、3の例では年齢が0歳の時を基準としていましたが、20歳の時を基準とする場合には、 $f(Z;\beta)-f(20;\beta)$ といった形で共変量である年齢に20を代入した式を対数ハザード比の式から減算します。

ハザード比

$$h(t;Z)/h(t;20)=h_0(t)\cdot\exp\big(f(Z;\beta)\big)/h_0(t)\cdot\exp\big(f(20;\beta)\big)$$
$$=\exp\big(f(Z;\beta)\big)/\exp\big(f(20;\beta)\big)$$
$$\Leftrightarrow$$

対数ハザード比

$$\ln(h(t;Z)/h(t;20))=f(Z;\beta)-f(20;\beta)$$

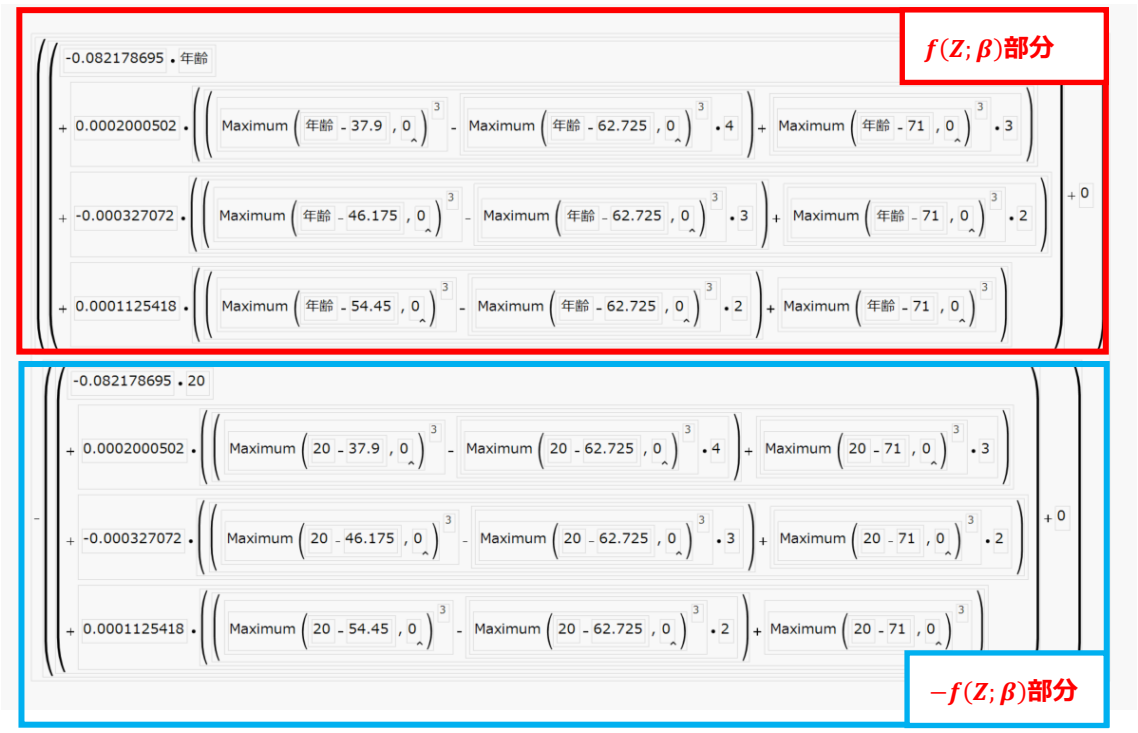
JMPで基準を変更する場合にはいくつかステップを踏まないといけません。＜操作:モデルのあてはめ＞を実施したうえで、次の操作を実施します。

＜操作:$f(Z;\beta)-f(20;\beta)$ の作成＞

1. データテーブルで[列]>[列の新規作成]を選びます。
2. 「列名」に「基準_年齢20」と入力し(列名は任意のもので構いません)、[OK]をクリックします。
3. データテーブルで、「線形予測子」列の表頭列名を右クリックして[列プロパティのコピー]を選びます。
4. データテーブルで、「基準_年齢20」の表頭列名を右クリックして[列プロパティの貼り付け]を選びます。
5. 同様に「基準_年齢20」の表頭列名を右クリックして[計算式]を選びます***。
6. 計算式エディタが開きますので、右側に表示される計算式をすべて選択して、右クリックから[コピー]を

選びます。

7. 同様に計算式がすべて選択された状態でエディタ上部の引き算ボタン「—」をクリックします。
8. 計算式エディタで選択していた部分の右側(ウィンドウによっては下側)に引き算記号とその隣に枠が表示されます。
9. 表示された枠を選び、右クリックから[貼り付け]を選びます。これにより$f(Z;\beta)-f(20;\beta)$が作成されました。
10. 計算式エディタで貼り付けた部分、すなわち$f(20;\beta)$で「年齢」と書かれた部分をすべて”20”に変更します。
11. [OK]をクリックします。



[グラフビルダー]でグラフを描く際には＜操作:[グラフビルダー]による可視化＞の手順 2.にて[Y]を「基準_年齢 20」に置き替えてください。

***＜操作:$f(Z;\beta)-f(20;\beta)$ の作成＞手順 6.以降については次の流れでも代替可能です。この方法は JMP のスクリプト言語(JSL)に詳しい方にお勧めです。

6'. 計算式エディタが開きますので、右側に表示される計算式をすべて選択してエディタ上部の引き算ボタン「—」をクリックします。
7'. 計算式エディタで選択していた部分の右側(ウィンドウによっては下側)に引き算記号とその隣に枠が表示されます。
8'. 表示された枠を選んでダブルクリックして編集可能な状態にします。
9'. 次の内容をコピー&ペーストします。

Eval(Substitute(:線形予測子 << get formula, Expr(:年齢), 20))

ここで、緑網掛け部分は＜操作:モデルのあてはめ＞手順5.で節点スプライン効果とした列の名前、黄色網掛け部分は基準とする値です。必要に応じて変更します。

10’. [OK]をクリックします。

参考資料

- Heinzl, H. and Kaider, A. (1997), Gaining More Flexibility in Cox Proportional Hazards Regression Models with Cubic Spline Functions, Computer Methods and Programs in Biomedicine, 54, 201-208.
- 一般化回帰モデル (jmp.com)
  https://www.jmp.com/support/help/ja/16.2/index.shtml#page/jmp/generalized-regression-models.shtml#
- 節点スプライン効果 (jmp.com)
  https://www.jmp.com/support/help/ja/16.2/index.shtml#page/jmp/knotted-spline-effect.shtml