

streg - パラメトリックモデル 【評価版】

streg は最尤法を用いたパラメトリックな生存時間分析機能を提供します。

1. パラメトリックモデル	
2. 基本的な用例	Example 1
	Example 2
	Example 3
3. 複数 failure 事象	Example 4
	Example 5
4. モデルの選択	Example 6
5. 補助パラメータのパラメータ化	Example 7
6. 層化推定	Example 8
	Example 9
7. 非共用 frailty モデル	Example 10
8. 共用 frailty モデル	Example 11



Stata14 で新たにサポートされた xtstreg, mestreg コマンドについてはそれぞれ [XT] xtstreg, [ME] mestreg をご参照ください。

1. パラメトリックモデル

パラメトリックモデルの場合には、ハザード関数、もしくは誤差分布の形状について関数形を明示した上で推定を行うこととなります。生存時間をパラメータ化する流儀にはいくつかありますが、streg コマンドでは PH モデル (proportional hazards model) と AFT モデル (accelerated failure-time model) の 2 種類に対応しています。

(1) PH モデル

この流儀ではハザード関数を

$$h(t) = h_0(t)g(\mathbf{x}) \quad (1)$$

のようにモデル化します。Cox 比例ハザードモデルの場合、 $h_0(t)$ については何も規定せずとも推定が行えたわけですが、パラメトリックモデルの場合には $h_0(t)$ について関数形を規定します。streg では次の 3 種類の PH モデルがサポートされています。

■ 指数モデル

モデル式 (1) において

$$h_0(t) = 1, \quad g(\mathbf{x}) = \exp(\beta' \mathbf{x}) \quad (2)$$

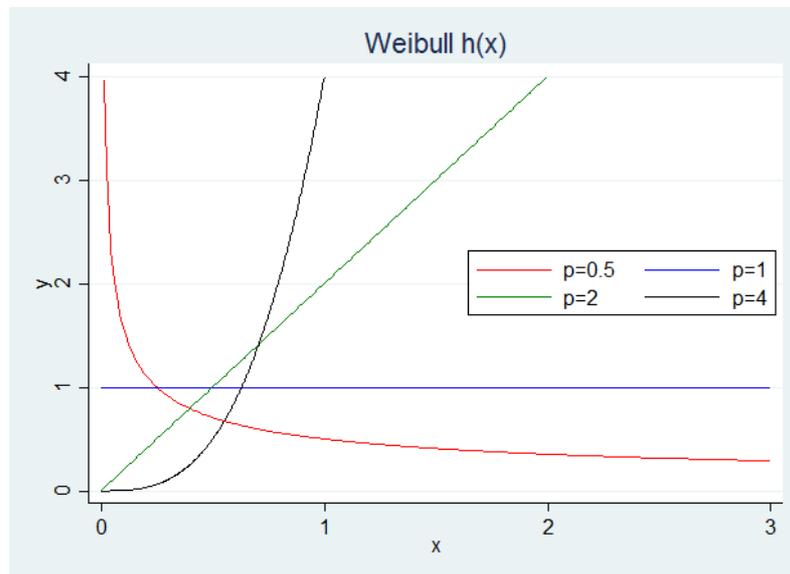
とモデル化するもので、Weibull モデルの $p = 1$ の場合に相当します。

■ Weibull モデル

モデル式 (1) において

$$h_0(t) = pt^{p-1}, \quad g(\mathbf{x}) = \exp(\beta' \mathbf{x}) \quad (3)$$

とモデル化するもので、補助パラメータ p も推定対象となります。ハザード関数の関数形状は次のようになります。

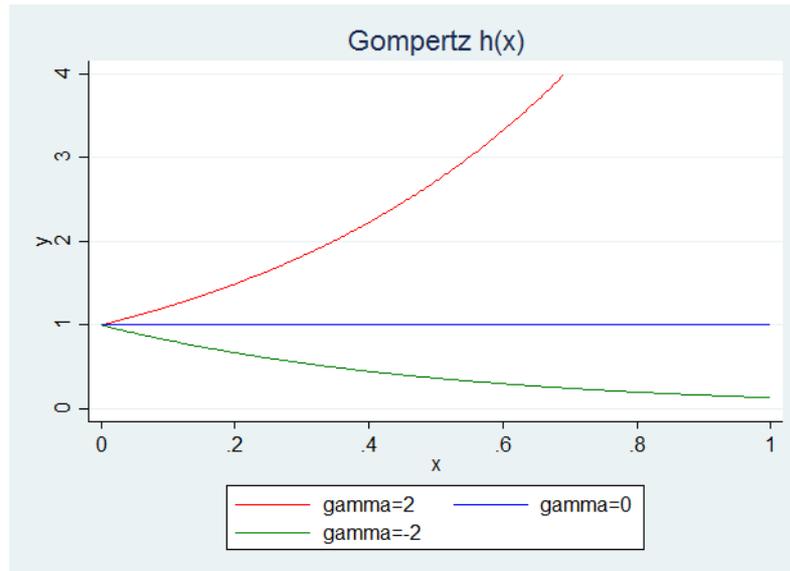


■ Gompertz モデル

モデル式 (1) において

$$h_0(t) = \exp(\gamma t), \quad g(x) = \exp(\beta'x) \quad (4)$$

とモデル化するもので、補助パラメータ γ も推定対象となります。ハザード関数の関数形状は次のようになります。



(2) AFT モデル

評価版では割愛しています。

2. 基本的な用例

▷ Example 1: Weibull – PH モデル

[ST] `streg` の Example 1 には Example データセット `kva.dta` を用いた用例が紹介されています。

```
. use http://www.stata-press.com/data/r14/kva.dta *1
(Generator experiment)
```

このデータセット中には発電機の耐久試験の結果が記録されています。

```
. list failtime load bearings *2
```

*1 メニュー操作 : File ▷ Example Datasets ▷ Stata 14 manual datasets と操作、Survival Analysis Reference Manual [ST] の `streg` の項よりダウンロードする。

*2 メニュー操作 : Data ▷ Describe data ▷ List data

	failtime	load	bearings
1.	100	15	0
2.	140	15	1
3.	97	20	0
4.	122	20	1
5.	84	25	0
6.	100	25	1
7.	54	30	0
8.	52	30	1
9.	40	35	0
10.	55	35	1
11.	22	40	0
12.	30	40	1

load は過負荷の量を、bearings は新しいタイプのベアリングを装着していたか否かを表す変数であり、failtime が故障が発生するまでの経過時間を意味しています。このデータセットは failtime を時間変数とする形で既に stset が済んでいます。このデータセットに対して Weibull(PH) モデルを前提とした形で streg を実行してみます。

- Statistics ▸ Survival analysis ▸ Regression models ▸ Parametric survival models と操作
- Model タブ: Independent variables: load bearings
Survival distribution: Weibull

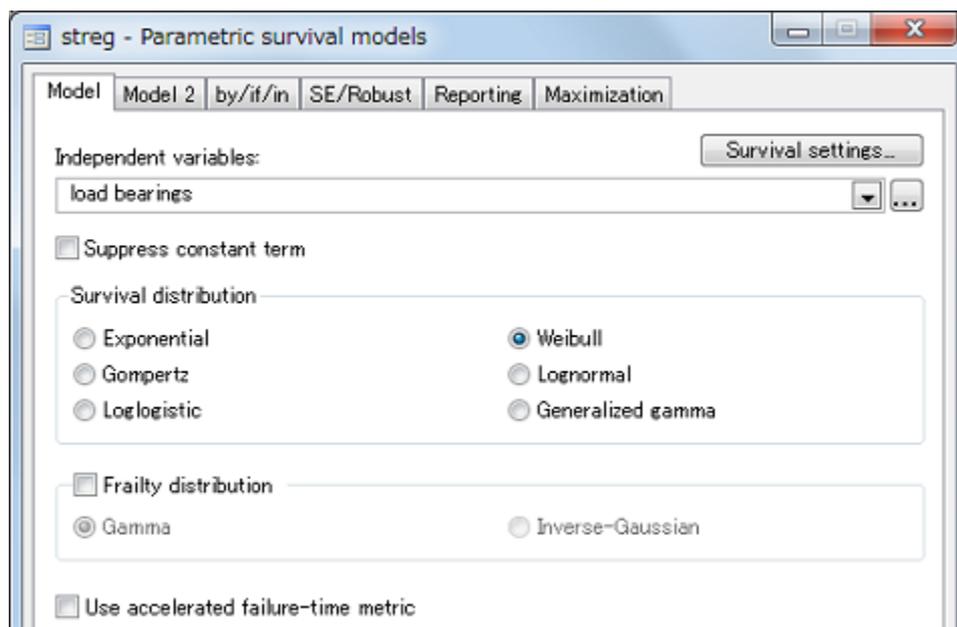


図1 streg ダイアログ- Model タブ

```

. streg load bearings, dist(weibull)

      failure _d:  1 (meaning all fail)
analysis time _t:  failtime

Fitting constant-only model:

Iteration 0:  log likelihood = -13.666193
Iteration 1:  log likelihood = -9.7427276
Iteration 2:  log likelihood = -9.4421169
Iteration 3:  log likelihood = -9.4408287
Iteration 4:  log likelihood = -9.4408286

Fitting full model:

Iteration 0:  log likelihood = -9.4408286
Iteration 1:  log likelihood = -2.078323
Iteration 2:  log likelihood =  5.2226016
Iteration 3:  log likelihood =  5.6745808
Iteration 4:  log likelihood =  5.6934031
Iteration 5:  log likelihood =  5.6934189
Iteration 6:  log likelihood =  5.6934189

Weibull regression -- log relative-hazard form

No. of subjects =          12          Number of obs   =          12
No. of failures =          12
Time at risk   =          896
Log likelihood =  5.6934189          LR chi2(2)       =          30.27
                                          Prob > chi2    =          0.0000

+-----+-----+-----+-----+-----+-----+
      _t | Haz. Ratio | Std. Err. | z | P>|z| | [95% Conf. Interval]
+-----+-----+-----+-----+-----+-----+
      load | 1.599315 | .1883807 | 3.99 | 0.000 | 1.269616 | 2.014631
bearings | .1887995 | .1312109 | -2.40 | 0.016 | .0483546 | .7371644
   _cons | 2.51e-20 | 2.66e-19 | -4.26 | 0.000 | 2.35e-29 | 2.68e-11
+-----+-----+-----+-----+-----+-----+
   /ln_p | 2.051552 | .2317074 | 8.85 | 0.000 | 1.597414 | 2.505691
+-----+-----+-----+-----+-----+-----+
      p | 7.779969 | 1.802677 | | | | 4.940241 | 12.25202
   1/p | .1285352 | .0297826 | | | | .0816192 | .2024193
+-----+-----+-----+-----+-----+-----+

```

上の操作の場合、Model タブ上で Use accelerated failure-time metric というオプションを選択していないため、PH メトリックを前提とした形で推定が行われています。[ST] `stcox` (*mwp-023*) の Example 1 では同じモデルを `stcox` によってフィットさせたわけですが、推定結果を対比させてみると次のようになります。

	ハザード比推定値	
	load	bearings
stcox	1.53	0.06
streg	1.60	0.19

streg の場合、ベースラインハザード関数の推定もフィットされており、Weibull 分布の形状パラメータが $\hat{p} = 7.78$ と推定されています。この結果からすると、100 時間経過後のベアリングは 10 時間経過後のベアリングに比べ、1,000,000 倍 $((100/10)^{7.78-1})$ 以上こわれやすいことがわかります。 <

▷ Example 2: 係数値出力

評価版では割愛しています。

▷ Example 3: Weibull – AFT モデル

評価版では割愛しています。

3. 複数 failure 事象

▷ Example 4

評価版では割愛しています。

▷ Example 5

評価版では割愛しています。

4. モデルの選択

▷ Example 6

評価版では割愛しています。

5. 補助パラメータのパラメータ化

表 1 に示されているように、ほとんどの分布形には補助パラメータ (ancillary parameters) が伴います。Weibull 分布の場合の p がその一例です。通常これらの補助パラメータは定数値として推定が行われますが、共変量に依存した形にすることも可能です。

▷ Example 7

評価版では割愛しています。

6. 層化推定

層化推定 (stratified estimation) の場合、ハザード関数は層 (グループ) ごとに異なることが許容されます。このような形の推定を行うには通常 `strata()` オプションが使用されますが、より制限された形のものであれば `ancillary()` オプションを利用する方法もあります。 `strata()` オプションを使用した場合にはモデル式中の切片項が層ごとに異なる値を取るばかりでなく、補助パラメータの値も層ごとに異なることが仮定されます。

▷ Example 8

評価版では割愛しています。

▷ Example 9: 部分的な層化

評価版では割愛しています。

7. 非共用 frailty モデル

評価版では割愛しています。

▷ Example 10

評価版では割愛しています。

8. 共用 frailty モデル

共用 frailty モデルの場合、frailty を入れ込む対象は個人ではなくグループとなります。すなわちグループ i に属する観測データ j のハザードは

$$h_{ij}(t|\alpha_i) = \alpha_i h_{ij}(t) \quad (5)$$

のように表現されます。ただし $i = 1, \dots, n$, $j = 1, \dots, n_i$ とし、 $h_{ij}(t)$ は $h(t|x_{ij})$ — 共変量 x_{ij} が与えられたときの個体レベルのハザード — を意味するものとします。

共用 frailty モデルは異質性がグループに固有のものであると考えられる場合に適したモデルであると言えます。グループに属する個体がすべて同一の frailty α_i を共有することによって、グループ内の相関がモデル化されることになるわけです。

▷ Example 11

評価版では割愛しています。

■