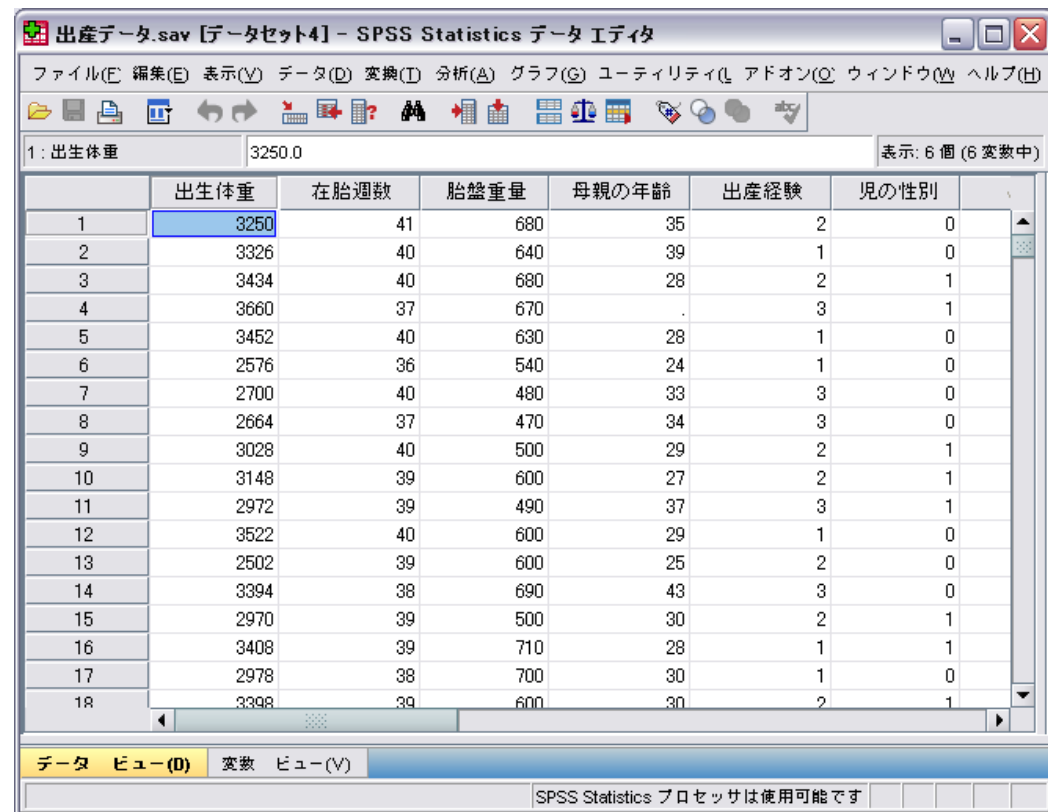


SPSSによるHosmer- Lemeshow検定について

データ例

- データは、http://www.hs.hirosaki-u.ac.jp/~pteiki/research/stat2/data_set.zipからダウンロードしたデータのうち、出産データ.savである



The screenshot shows the SPSS Statistics Data Editor window titled "出産データ.sav [データセット4] - SPSS Statistics データ エディタ". The window displays a data grid with 18 rows and 7 columns. The columns are labeled: 出生体重 (Birth Weight), 在胎週数 (Gestational Week), 胎盤重量 (Placental Weight), 母親の年齢 (Mother's Age), 出産経験 (Parity), and 児の性別 (Child's Sex). The first row is highlighted, showing a birth weight of 3250.0. The status bar at the bottom indicates "SPSS Statistics プロセッサは使用可能です".

	出生体重	在胎週数	胎盤重量	母親の年齢	出産経験	児の性別
1	3250	41	680	35	2	0
2	3326	40	640	39	1	0
3	3434	40	680	28	2	1
4	3660	37	670	.	3	1
5	3452	40	630	28	1	0
6	2576	36	540	24	1	0
7	2700	40	480	33	3	0
8	2664	37	470	34	3	0
9	3028	40	500	29	2	1
10	3148	39	600	27	2	1
11	2972	39	490	37	3	1
12	3522	40	600	29	1	0
13	2502	39	600	25	2	0
14	3394	38	690	43	3	0
15	2970	39	500	30	2	1
16	3408	39	710	28	1	1
17	2978	38	700	30	1	0
18	3398	39	600	30	2	1

SPSS12.0.1Jの結果

Hosmer と Lemeshow の検定

ステップ	カイ2乗	自由度	有意確率
1	5.193	8	.737
2	6.012	8	.646

Hosmer と Lemeshow の検定の分割表

		児の性別 = 女		児の性別 = 男		合計
		観測値	期待値	観測値	期待値	
ステップ1	1	83	87.843	60	55.157	143
	2	82	79.297	60	62.703	142
	3	86	76.610	58	67.390	144
	4	74	73.050	69	69.950	143
	5	68	70.174	74	71.826	142
	6	66	67.483	75	73.517	141
	7	70	66.046	73	76.954	143
	8	62	62.670	79	78.330	141
	9	58	60.008	84	81.992	142
	10	46	51.819	93	87.181	139
ステップ2	1	102	95.231	40	46.769	142
	2	88	85.176	54	56.824	142
	3	76	80.153	67	62.847	143
	4	65	74.957	77	67.043	142
	5	71	71.444	72	71.556	143
	6	65	67.392	77	74.608	142
	7	64	63.555	78	78.445	142
	8	65	59.703	77	82.297	142
	9	55	54.027	87	87.973	142
	10	44	43.362	96	96.638	140

- ロジスティック回帰分析の[オプション]ダイアログで, [表示]を[最後のステップ]にした結果, 上記の表が出力される.
- 12.0J以前では(13~15では確認していない), 最後のステップを選んでも, 全てのステップが出力される.
- この例で, 最終モデルのステップ2は独立変数が**出生体重**と**在胎週数**である. だとすると, Hosmer-Lemeshow検定の結果は, ステップ2の**p=0.646**である.

SPSS16.0J & 17.0Jの結果

Hosmer と Lemeshow の検定

ステップ	カイ2乗	自由度	有意確率
3	5.193	8	.737

Hosmer と Lemeshow の検定の分割表

		児の性別 = 女		児の性別 = 男		合計 (ピボットテーブル)
		観測	期待	観測	期待	
ステップ 3	1	83	87.843	60	55.157	143
	2	82	79.297	60	62.703	142
	3	86	76.610	58	67.390	144
	4	74	73.050	69	69.950	143
	5	68	70.174	74	71.826	142
	6	66	67.483	75	73.517	141
	7	70	66.046	73	76.954	143
	8	62	62.670	79	78.330	141
	9	58	60.008	84	81.992	142
	10	46	51.819	93	87.181	139

- 16以降では、最終ステップのみの結果が出力されるようである...
- 先の例と同じ手順で行っている。
- しかし、有意確率は $p=0.737$ で、先の例のステップ1(出生体重のみのモデル)と同じ値になっている!
- しかもステップ3となっている!

もう一度SPSS16.0J & 17.0Jの結果

Hosmer と Lemeshow の検定

ステップ	カイ 2 乗	自由度	有意確率
1	5.193	8	.737
2	6.012	8	.646

Hosmer と Lemeshow の検定の分割表

		児の性別=女		児の性別=男		合計 (ピボットテーブル)
		観測	期待	観測	期待	
ステップ 1	1	83	87.843	60	55.157	143
	2	82	79.297	60	62.703	142
	3	86	76.610	58	67.390	144
	4	74	73.050	69	69.950	143
	5	68	70.174	74	71.826	142
	6	66	67.483	75	73.517	141
	7	70	66.046	73	76.954	143
	8	62	62.670	79	78.330	141
	9	58	60.008	84	81.992	142
	10	46	51.819	93	87.181	139
ステップ 2	1	102	95.231	40	46.769	142
	2	88	85.176	54	56.824	142
	3	76	80.153	67	62.847	143
	4	65	74.957	77	67.043	142
	5	71	71.444	72	71.556	143
	6	65	67.392	77	74.608	142
	7	64	63.555	78	78.445	142
	8	65	59.703	77	82.297	142
	9	55	54.027	87	87.973	142
	10	44	43.362	96	96.638	140

- [オプション]ダイアログで,
[表示]を[ステップごと]にして出力すると...
- 今度は, ステップ2が出力され, $p=0.646$ となる
- 他のデータ例で照合してみたが, どうも, SPSS16以降(13~15では未確認)では, **最終ステップのみの出力**とすると値がおかしくなるようである

結論

- 現状では、ロジスティック回帰分析を行うときは[ステップごと]にした方が良いでしょう。
- Hosmer-Lemeshow 検定以外は、同じ値で出力される。

