ISSN : 2460 – 7797
e-ISSN : 2614-8234
Website : jurnal.umj.ac.id/index.php/fbc
Email : fibonacci@umj.ac.id

FIBONACCI
Jurnal Pendidikan Matematika dan Matematika

MEMAHAMI PENGGUNAAN REGRESI PADA DATA RESPON MULTINOMIAL UNTUK PENELITIAN SOSIAL DAN KEPENDIDIKAN

Halimatus Sa'diyah [1)*] dan Riza Yuli Rusdiana [2)]
[1,2)] Lab. Biometrika, Fakultas Pertanian, Universitas Jember

[*] *sadiyah@unej.ac.id*

Abstrak

Model logit multinomial digunakan untuk memodelkan sifat hubungan antara peubah respon politomus dan peubah penjelas. Ada dua model logit multinomial untuk peubah respon politomus yang strukturnya tak berurut: model logit terampat dan model logit bersyarat. Kedua model mempunyai struktur serupa, $g[P(Y = j)] = x^T \beta_j$, *j = 1, …, k, tetapi berbeda dalam karakteristik peubah penjelasnya. Logit terampat menggunakan karakteristik dari individu (subyek) sebagai peubah penjelas, sedang logit bersyarat menggunakan karakteristik dari pilihan individu. Tulisan ini ingin menyajikan penggunaan keduanya dalam model regresi yang sering dibutuhkan dalam penelitian-penelitian sosial dan kependidikan. Ilustrasi melalui data hipotetik digunakan untuk memperjelas kebutuhan, kegunaan, metode analisis data sampai pada interpretasi hasil model regrgesi. Sajian komputasi yang ringkas dilakukan melalui dua software yang populer yaitu SAS untuk model regresi generalized-logit. Sedangkan untuk model regresi logit bersyarat dapat dilakukan dengan SAS dan SPSS. Baik SAS dan SPSS menyajikan hasil analsis regresi yang sama untuk model regresi logit bersyarat.*

Kata Kunci: *Regresi, Nominal, Logit, Bersyarat.*

PENDAHULUAN

Setiap model statistika didasarkan pada: yang pertama, gagasan bahwa peubah acak yang diselidiki mempunyai struktur yang dapat menjelaskan nilai-nilai yang diperoleh, dan memprediksi nilai-nilai untuk studi berikutnya. Keacakan justru menjadi sesuatu yang sangat penting dalam prediksi statistika ini. Bahkan Hadi (2021) menyebutkan keacakan sebagai sesuatu yang wajib ada dalam apa yang disebut sebagai "gambar besar" (*big picture*) dalam Statistika Intuitif (Hadi, 2021).

Yang kedua, adalah postulat bahwa peubah acak yang diselidiki dapat diungkapkan atau diekspresikan ke dalam suku-suku yang disebut komponen struktur. Jika suku-suku ini mempunyai nilai tetap

DOI: https://dx.doi.org/10.24853/fbc.7.2.109-126

(meskipun tidak diketahui), maka ia disebut komponen sistematik dan jika bernilai acak disebut komponen acak. Struktur ini dianggap sebagai deskripsi populasi yang nilai-nilainya diberikan oleh nilai-nilai dari contoh acak (*random sample*). Dengan perkataan lain, setiap model statistika selalu terdiri atas dua komponen, yaitu (1) komponen sistematik yang menjelaskan perilaku data, dan (2) komponen acak yang menjelaskan keragaman data (McCullagh dan Nelder, 1989).

Model logit multinomial digunakan untuk memodelkan sifat hubungan antara peubah respon politomus dan segugus peubah penjelas (*covariate, explanatory variable, regressor variable*). Model-model respon politomus ini dapat dikelompokkan ke dalam dua kelas model yang berbeda bergantung pada apakah peubah respon yang diselidiki mempunyai struktur terurut (*ordered*) ataukah tak berurut (*unordered*). Naskah ini hanya berkenaan dengan model logit multinomial yang peubah responnya mempunyai struktur tak berurut atau yang berskala nominal.

Ada dua model logit multinomial untuk peubah respon nominal, yaitu model logit terampat (generalized logit model) dan model logit bersyarat (conditional logit model). Kedua model ini mempunyai struktur model serupa,

$$g[P(Y=j)] = x^T\beta_j, \quad j = 1, \ldots, k,$$

tetapi berbeda dalam karakteristik peubah penjelas yang digunakan. Disini, g adalah model sebaran peluang. Peubah penjelas model logit terampat menggunakan karakteristik dari individu (subjek), sedang model logit bersyarat menggunakan karakteristik dari pilihan individu. Untuk memperjelas pengertian tentang kedua model dan perbedaan antara karakteristik individu dan karakteristik pilihan individu, marilah kita mencermati tiga ilustrasi berikut:

Ilustrasi 1. Simaklah data dalam List 1 (dalam format tahapan DATA SAS) tentang pilihan jenis angkutan (buatan). Anggaplah sejumlah subyek (responden) diminta untuk memilih jenis angkutan yang digunakan: pesawat, kereta-bisnis, ataukah bis-cepat. Pilihan didasarkan atas pertimbangan biaya (dalam hal ini satuan waktu yang dihabiskan dalam perjalanan mulai dari asal sampai tujuan).

List 1. Pilihan jenis angkutan

```
data angkutan;
  input wpesawat wkereta wbis
umur preferensi $ frek;
  datalines;
4.5 10.0 10.5 32 pesawat 1
6.0  4.5  5.5 41 bis     1
...
4.0  1.5  2.0 22 kereta  1
;
```

Peubah `wpesawat`, `wkereta`, dan `wbis` merupakan peubah-peubah yang menjadi pertimbangan preferensi atau pilihan subyek, sedang umur merupakan karakteristik subyek (responden).

Model logit terampat digunakan untuk menyelidiki sifat hubungan antara preferensi angkutan dan umur, sedang model logit bersyarat digunakan untuk meyelidiki bagaimana waktu tempuh mempengaruhi preferensi responden. Untuk menyelidiki bagaimana kedua jenis peubah penjelas (waktu dan umur) mempengaruhi preferensi jenis angkutan, kita memerlukan model campuran (*mixed model*) yang menggabungkan kedua jenis peubah penjelas.

Ilustrasi 2. Hasil riset Ries dan Smith tahun 1963 (Agresti, 1996) dalam List 2 berkenaan dengan preferensi subyek terhadap merek deterjen M dan X yang dihubungkan dengan tiga peubah penjelas: kelembutan (*softness*) air cucian (lembut, sedang, atau kasar), temperatur air (tinggi/rendah), dan subyek (responden) pernah menggunakan merek M (Ya/Tidak).

List 2. Preferensi merek deterjen M dan X.

```
data deterjen;
  input pernahM $ suhu $ cairan
$ preferensi $ frek;
  datalines;
ya     rendah  lembut  M  49
ya     rendah  sedang  M  55
ya     rendah  kasar   M  52
...
tidak  tinggi  lembut  X  29
tidak  tinggi  sedang  X  33
tidak  tinggi  kasar   X  42
;
```

Dalam kasus ini, peubah-peubah pernahM merupakan karakteristik individu, sedang suhu dan cairan adalah karakteristik dari pilihan individu.

Ilustrasi 3. Data dalam List 3 merupakan hasil survei General Social Survey pada tahun 1991 (Agresti 1996, hal 169, urutan pertanyaan diubah) untuk tiga pertanyaan yang diajukan kepada kaum kulit putih. Ketiga pertanyaan (A, B, dan C) tersebut adalah

A. Jika partai Anda mencalonkan seorang Presiden yang berkulit hitam (baca: negro), akankah Anda memilihnya jika ia pantas (*qualified*) untuk jenis pekerjaan (jabatan) itu?
B. Sukakah Anda (responden kulit putih) mengantarkan anak-anak sekolah negro dan kulit putih dari satu tempat ke tempat lain?
C. Selama beberapa tahun terakhir ini, adakah di antara anggota keluarga Anda (termasuk Anda) mengajak teman negro untuk makan malam bersama dengan keluarga Anda?

Skala respon untuk setiap jenis pertanyaan adalah 1=Ya, 2=Tidak, dan 3=Tidak tahu. Untuk kasus kita disini, jawaban yang diinginkan (mungkin) adalah apakah ada hubungan antara peubah respon A dengan peubah penjelas B dan C. Dari jenis pertanyaan tersebut, jelas bahwa B dan C bukan karakteristik pilihan individu, tetapi karakteristik individu.

List 3. Preferensi kaum kulit putih terhadap calon Presiden kulit hitam (negro).

```
data presiden;
  input A B C frek;
  datalines;
1 1 1  41
1 1 2  65
1 1 3   0
...
3 3 3   1
;
```

Tujuan dari naskah ini adalah mempelajari penerapan model-model logit multinomial pada peubah respon tak berurut (*unordered*) atau skala nominal..

METODE PENELITIAN

<u>Sebaran Multinomial</u>

Diberikan contoh acak berukuran n. Jika hasil setiap pengamatan bebas hanya termasuk ke dalam satu dan hanya satu dari J kategori,

$$Y_j \in \{Y_1, Y_2, \ldots, Y_J\}, \text{ dimana } j = 1, \ldots, J,$$
$$\text{dan } \sum_{j=1}^{J} Y_j = n,$$

maka Yj dikatakan mengikuti sebaran peluang multinomial. Peluang contoh yang berukuran n mengandung nj kategori ke-j adalah

$$P(n_1, n_2, \dots, n_J) = \frac{n!}{n_1! n_2! \dots n_J!} p_1^{n_1} p_2^{n_2} \dots . p_J^{n_J}$$

dengan $\sum_{j=1}^{J} p_j = 1$.

Suatu transformasi logit untuk peluang multinomial $p_1, p_2, \dots, p_j$ ke gugus parameter $\pi_1, \pi_2, \dots, \pi_j$ yang didefinisikan sebagai

$$\pi_j = log\left(\frac{p_j}{p_J}\right) \text{ untuk } j = 1, \dots, J.$$

disebut logit multinomial. Catatlah bahwa $\pi_j = 0$, sehingga ada $j - 1$ parameter logit. Definisi ini menggunakan kategori terakhir (*j*) sebagai kategori-landasan (*baseline-categorical*) atau kategori-acuan (*reference-categorical*). Pemilihan kategori-landasan ini tidak khas. Sembarang kategori *j* (*j = 1, ..., j*) dapat dipilih sebagai kategori-landasan. Misalnya, jika logit multinomial didefinisikan sebagai

$$\pi_j^* = log\left(\frac{p_j}{p_1}\right)$$

sehingga $\pi_1^* = 0$, maka π_j^* hanyalah reparameterisasi π_j, karena

$$\pi_j^* = \pi_j + log\left(\frac{p_J}{p_1}\right).$$

Logit multinomial memainkan peranan penting dalam analisis data respon multi-kategorik seperti halnya logit binomial dalam tabel dua-arah. Data hasil pengamatan untuk kasus di atas biasanya disajikan dalam bentuk matriks data/peluang atau tabel kontingensi (Gambar 1). Dalam konteks ini (Gambar 1), sel-sel dalam tabel yang tidak mempunyai kategori biasanya diperlakukan sebagai contoh acak Poisson bebas. Jika J = 2 (ada dua kategori), maka baris sel dalam tabel kontingensi yang berukuran I × J mempunyai sebaran binomial; dan respon dari peubah acak Y1 dan Y2 disebut respon biner atau respon kuantal. Jika setiap baris tabel merupakan contoh bebas, maka rancangannya disebut rancangan binomial bebas. Untuk J ≥ 3, setiap baris sel dalam tabel kontingensi mempunyai sebaran multinomial; dan respon dari peubah acak Y_1, Y_2, …, Y_J disebut respon multikategorik (*multicategorical response*), juga disebut respon polikotomus (*polychotomous response*) atau respon politomus (*polytomous response*). Serupa dengan rancangan binomial bebas, jika setiap baris tabel saling bebas, maka rancangan disebut rancangan multinomial bebas (Agresti 1996).

	Kategori Respon				
$Contoh$	1	2	$\cdots$	J	$Total$
1	n_{11}	n_{12}	$\cdots$	n_{1J}	$n_{1\,.}$
2	n_{21}	n_{22}	$\cdots$	n_{2J}	$n_{2\,.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
I	n_{I1}	n_{J2}	$\cdots$	n_{IJ}	$n_{I\,.}$
n_{11}	$n_{.\,1}$	$n_{.\,2}$	$\cdots$	$n_{.\,J}$	n

	Peluang Respon						
Contoh	1	2	$\cdots$	j	$\cdots$	J	Total
1	π_{11}	π_{12}	$\cdots$	π_{1j}	$\cdots$	π_{1J}	$\pi_{1\,.}$
2	π_{21}	π_{22}	$\cdots$	π_{2j}	$\cdots$	π_{2J}	$\pi_{2\,.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	π_{i1}	π_{i2}	$\cdots$	$\pi_{ij} = \frac{n_{ij}}{n}$	$\cdots$	π_{iJ}	$\pi_{i\,.} = \sum_{j=1}^{J} \pi_{ij}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$	$\vdots$
I	π_{I1}	π_{J2}	$\cdots$	π_{Ij}	$\cdots$	π_{IJ}	$\pi_{I\,.}$
	$\pi_{.\,1}$	$\pi_{.\,2}$	$\cdots$	$\pi_{.\,j} = \sum_{i=1}^{I} \pi_{ij}$		$\pi_{.\,J}$	1

Gambar 1. Matriks data atau tabel kontingensi hasil hitungan (*count data*) dan nilai peluangnya

Merujuk pada Gambar 1, kategori baris tabel mungkin tunggal, mungkin majemuk. Setiap peubah kategori dalam baris matriks atau tabel kontingensi, dalam konteks bahasan ini, disebut sebagai peubah penjelas (*covariate, explanatory variable*).

Peluang Respon Multinomial

Misalkan responden dihadapkan pada J alternatif yang tersedia. Andaikan π_{jk} (j = 1, …, n, k = 1, …, J) menyatakan peluang bahwa responden j memilih alternatif k, X_j menyatakan karakteristik individu j, dan Z_{jk} menyatakan karakteristik dari alternatif k yang menjadi pilihan individu j. Dalam ilustrasi 1, X_j adalah umur dan setiap Z_{jk} adalah waktu perjalanan. Dalam model logit terampat, setiap subyek/individu dianggap sebagai satuan analisis (*analysis unit*) dan karakteristik individu digunakan sebagai peubah penjelas. Sebagai peubah penjelas, karakteristik individu dianggap konstan terhadap alternatif yang tersedia.

Misalkan, setiap *J* pilihan jenis angkutan dalam ilustrasi 1, $X_j = (1 \text{ umur})^T$, dan untuk responden pertama $X_1 = (1\ 32)^T$. Peluang bahwa individu *j* memilih alternatif *k* adalah

$$\pi_{jk} = \frac{exp(\beta_k^T x_j)}{\sum_{m=1}^{n} exp(\beta_m^T x_j)} = \frac{1}{\sum_{m=1}^{n} exp[(\beta_m^T - \beta_k^T) x_j]} \qquad [1]$$

dimana β1, β1, …, βn, adalah vektor parameter regresi berukuran n (tidak diketahui), meskipun Xj konstan untuk seluruh alternatif. Karena $\sum_{j=1}^{n} \pi_{jk} = 1$, maka n gugus dari parameter ini tidak khas.

Jika gugus terakhir dipilih sama dengan 0 ($\beta_n = 0$), maka β_k mewakili pengaruh peubah *X* terhadap peluang memilih alternatif *k* relatif terhadap alternatif terakhir. Dengan memilih $\beta_n = 0$, ada $n - 1$ gugus koefisien regresi atau persamaan regresi logit. Model logit terampat oleh Agresti (1990 & 1996) disebut sebagai model logit berkategori-landasan (*baseline-categorical logit model*). Beberapa penulis lain mengatakan model logit yang mempunyai kategori acuan (*reference-categorical logit model*) (Aitkin *et. al*. 1989).

Pada software SAS, model logit terampat dikerjakan dengan PROC CATMOD, dan dalam SPSS ditetapkan dengan prosedur GENLOG. Beberapa software lain yang dapat digunakan untuk analisis model logit terampat adalah MINITAB 13.0, STATA, S-PLUS 2000, dan mungkin masih banyak yang lain. Model ini banyak digunakan dalam berbagai bidang, misalnya: pertanian, sosial, ekonomi, politik, dan hukum khususnya untuk evaluasi kebijakan. Dalam bidang keproyekan banyak digunakan untuk studi BME (*benefit, monitoring, and evaluation*) dan mengkaji informasi *benchmark*.

Dalam model logit bersyarat, peubah penjelas *Z* dianggap mempunyai nilai-nilai alternatif berbeda, dan dampak (*impact*) dari unit *Z* diasumsikan konstan terhadap alternatif yang lain. Dalam ilustrasi 1, untuk setiap *n* pilihan jenis angkutan $Z_{jk} = (\text{waktu})^T$. Untuk subyek pertama nilai $Z_{11} = (4.5)^T$, $Z_{12} = (10.0)^T$, dan $Z_{12} = (10.5)^T$. Peluang bahwa subyek *j* memilih alternatif *k* adalah

$$\pi_{jk} = \frac{exp(\theta^T Z_{jk})}{\sum_{m=1}^{n} exp(\theta^T Z_{jm})} = \frac{1}{\sum_{j=1}^{n} exp[\theta^T (Z_{jm} - Z_{jk})]} \qquad [2]$$

dimana θ adalah vektor koefisien regresi tunggal. Pengaruh peubah penjelas terhadap peluang alternatif yang dipilih berasal dari perbedaan antara nilai-nilai dari alternatif. Dengan perkataan lain, sifat dari Persamaan [2] ini adalah kepastian (*odds*) bahwa individu *j* memilih alternatif *k* relatif terhadap alternatif ℓ merupakan fungsi dari perbedaan antara Z_{jk} dan $Z_{j\ell}$ (karakteristik pilihan subyek, dan bukan karakteristik subyek):

$$log\left(\frac{\pi_{jk}}{\pi_{j\ell}}\right) = \beta\left(Z_{jk} - Z_{j\ell}\right).$$

Jadi kepastian suatu subyek j memilih alternatif k relatif terhadap alternatif ℓ tidak bergantung pada pilihan-pilihan (option) lain dalam gugus pilihan yang tersedia atau nilai-nilai yang lain pada peubah atribut. Oleh sebab itu, model dapat diungkapkan dalam bentuk model logit bersyarat yang pendugaannya didasarkan pada model resiko kesebandingan (Agresti, 1990).

Dalam SAS, model logit bersyarat dapat dikerjakan dengan PROC PHREG, dan dalam SPSS dapat dikerjakan dengan COXREG. Model ini didasarkan pada model resiko kesebandingan (*proportional hazard model*) yang banyak digunakan dalam riset kesehatan (*medical reasearch*), riset penyakit tanaman, atau analisis model-model yang mendasarkan pada analisis peluang-hidup (*survival analysis*). Dalam riset pemasaran digunakan untuk analisis preferensi merek dan merupakan alternatif analisis konjoin (Ying So dan Kuhfeld, 1995). Di bidang pertanian, digunakan untuk uji preferensi hama atau penyakit terhadap varietas tanaman.

Model logit campuran (*mixed logit model*) menggabungkan karakteristik subyek dan karakteristik dari pilihan subyek (Ying So dan Kuhfeld, 1995), mempunyai bentuk peluang

$$\pi_{jk} = \frac{exp\left(\beta_k^T X_j + \theta^T Z_{jk}\right)}{\sum_{m=1}^{n} exp\left(\beta_m^T X_j + \theta^T Z_{jm}\right)}$$

Kita tidak membahas model ini. Bagi yang tertarik, dapat membaca lebih lanjut pada kepustakaan naskah ini, terutama pada Ying So & Kuhfeld (1995).

HASIL DAN PEMBAHASAN

Bagian ini akan membahas penerapan model-model nominal dan mendiskusikannya dalam beberapa kegunaan pada penelitian sosial juga pada penggunaan software yang berbeda. Penerapan model regresi nominal diawali dengan model multinomial-logit dengan software SAS dan SPSS.

Aplikasi Logit Multinomial: Respon Nominal

Regresi Logit Terampat dalam SAS: PROC CATMOD

Model regresi multinomial, dalam SAS, dikerjakan dengan PROC CATMOD. Jika x_1 kuantitatif dan x_2 kualitatif, sintaks untuk model logit terampat dalam SAS adalah:

```
proc catmod;
   direct x1;
   response logits;
   model y=x1 x2;
run;
```

Pernyataan `DIRECT` menyatakan bahwa x_1 kontinu, dan ini harus diletakkan sebelum pernyataan `MODEL`. Pernyataan `RESPONSE` menspesifikkan fungsi peluang respon yang digunakan untuk memodelkan fungsi respon sebagai kombinasi linear dari unsur-unsur vektor parameter β.

Tipe `RESPONSE` yang digunakan dalam `PROC CATMOD` bergantung pada model yang digunakan:

`CLOGITS` (*cumulative logits*):

$$log\left(\frac{p_1}{1-p_1}\right),\ \ log\left(\frac{p_1+p_2}{1-p_1-p_2}\right),\ \ ....,$$
$$log\left(\frac{p_1+p_2+.....+p_J}{1-p_1-p_2-.....-p_J}\right)$$

`ALOGITS` (*adjacent logits*):

$$log\left(\frac{p_2}{p_1}\right),\ \ log\left(\frac{p_3}{p_2}\right),\ \ ....,\ \ log\left(\frac{p_{k+1}}{p_J}\right)$$

Tipe `RESPONSE` memiliki nilai *default* LOGITS (*generalized logits*) yang mempunyai bentuk (Guido, et al, 2006):

$$log\left(\frac{p_j}{p_J}\right) = x^T\beta_j, \text{ for } j = 1, ..., k.$$

Sebagai catatan, model logit terampat bergantung pada pendekatan asimptotik (*assymptotic approximations*), jadi diperlukan contoh yang cukup besar untuk mendukung fungsi sebaran normal asimptotik. Sebagai pedoman umum (SAS Technical Support, 2005), menyarankan sedikitnya ada 25–30 contoh untuk setiap fungsi respon. Misalnya, jika ada satu peubah bebas dengan J = 4 kategori, maka model logit terampat (ada tiga model) masing-masing memerlukan sedikitnya 75. Subyek juga perlu cukup tersebar ke seluruh sel-sel tabel sehingga kurang dari 20% fungsi respon yang mempunyai ukuran contoh efektif kurang dari 5. Jika setiap populasi ukurannya kurang dari 5 subyek dalam fungsi respon kategori pertama, maka lebih baik menggabungkan kategori ini ke kategori yang lain daripada memaksakan asumsi kenormalan asimptotik untuk fungsi respon. Jika peubah respon mempunyai skala ordinal, sebaiknya menggunakan rataan fungsi respon daripada meminta tiga logit terampat. Jika ada lebih dari satu peubah respon, dan Anda menspesifikkan rata-rata respon, maka ukuran contoh efektif 30 cukup untuk mendukung empat fungsi respon

Teladan disini menggunakan data coklat (McCullagh dan Nelder, 1989) tetapi beberapa frekuensinya dimodifikasi tanpa mengurangi total frekuensi. Ini dilakukan untuk menghindari sparseness yang menimbulkan masalah dalam pemodelan model logit terampat. Versi modifikasi juga digunakan dalam manual SAS (SAS Institute, 2000:435-438). Untuk menyederhanakan perhitungan, dalam teladan ini, kategori respon diubah dari 9 (aslinya ordinal, y) menjadi hanya 3 (dianggap nominal, ymin).

List 4. Teladan `PROC CATMOD` dengan data coklat

```
data coklat;
   input y x1 x2 x3 x4 frek;
   if y=1 or y=2 or y=3 then
ymin=1;  /* baris ini digunakan
untuk */
      else if y=4 or y=5 or y=6
then ymin=2; /*mengkonversi 9 skala
*/
           else ymin=3;
   /* ordinal menjadi 3 skala
   nominal*/
     *Kode peubah baru x (dalam
   pembahasan) sisipkan disini;
datalines;
   1 1 0 0 0  1
   2 1 0 0 0  1
   3 1 0 0 0  1
   4 1 0 0 0  5
   5 1 0 0 0  8
   6 1 0 0 0  8
   7 1 0 0 0 19
   8 1 0 0 0  8
   9 1 0 0 0  1
   1 0 1 0 0  6
   2 0 1 0 0  9
   3 0 1 0 0 12
   4 0 1 0 0 11
   5 0 1 0 0  7
   6 0 1 0 0  4
   7 0 1 0 0  1
   8 0 1 0 0  1
   9 0 1 0 0  1
   1 0 0 1 0  1
   2 0 0 1 0  1
   3 0 0 1 0  6
   4 0 0 1 0  8
   5 0 0 1 0 23
   6 0 0 1 0  7
   7 0 0 1 0  4
   8 0 0 1 0  1
   9 0 0 1 0  1
   1 0 0 0 1  1
   2 0 0 0 1  1
   3 0 0 0 1  1
   4 0 0 0 1  1
   5 0 0 0 1  1
   6 0 0 0 1  6
   7 0 0 0 1 14
   8 0 0 0 1 16
   9 0 0 0 1 11
   ;
   proc catmod data=coklat;
      direct x1-x4;
     weight frek;
     response logits;
      model ymin=x1-x4 / noiter
   freq;
   run;
*/Kode proc catmod dengan x kategorik (dalam
pembahasan) dapat sisipkan disini*/ quit;
```

Output dari program di atas adalah

```
                     The CATMOD  Procedure

Response            ymin       Response Levels    3
Weight Variable     frek       Populations        4
Data Set            COKLAT     Total Frequency  208
Frequency Missing   0          Observations      36


                     Population Profiles
                                                  Sample
    Sample     x1     x2     x3     x4             Size
    ------------------------------------------------------
        1      0      0      0      1                52
        2      0      0      1      0                52
        3      0      1      0      0                52
        4      1      0      0      0                52


                       Response Profiles
                       Response    ymin
                       -----------------
                           1       1
                           2       2
                           3       3


                        Response Frequencies
                           Response Number
        Sample         1          2          3
        ---------------------------------------
            1          3          8         41
            2          8         38          6
            3         27         22          3
            4          3         21         28


        Maximum Likelihood Analysis of Variance

Source                 DF    Chi-Square    Pr > ChiSq
------------------------------------------------------
Intercept               2         33.46        <.0001
x1                      2          7.79        0.0203
x2                      2         36.33        <.0001
x3                      2         37.07        <.0001
x4                      0*            .             .

Likelihood Ratio        0             .             .

NOTE: Effects marked with '*' contain one or more
      redundant or restricted parameters.


            Analysis of Maximum Likelihood Estimates

                                         Standard       Chi-
Effect       Parameter    Estimate        Error       Square    Pr > ChiSq
---------------------------------------------------------------------------
Intercept         1        -2.6150       0.5981        19.12        <.0001
                  2        -1.6341       0.3865        17.88        <.0001
x1                3         0.3814       0.8525         0.20        0.6546
                  4         1.3464       0.4824         7.79        0.0053
x2                5         4.8122       0.8532        31.81        <.0001
                  6         3.6266       0.7267        24.90        <.0001
x3                7         2.9026       0.8058        12.97        0.0003
                  8         3.4800       0.5851        35.37        <.0001
x4                9              .            .            .             .
                 10              .            .            .             .
```

Dari analisis diperoleh dua model regresi dugaan:

$$log\left(\frac{p_1}{p_3}\right) = -2.6150 + 0.3814X_1 + 4.8122X_2 + 2.9026X_3$$

$$log\left(\frac{p_2}{p_3}\right) = -1.6341 + 1.3464X_1 + 3.6266X_2 + 3.4800X_3$$

Model logit yang diduga pada setiap kombinasi matriks rancangan X adalah:

X_1	X_2	X_3	X_4	$log\left(\frac{p_1}{p_3}\right)$	$log\left(\frac{p_2}{p_3}\right)$
1	0	0	0	-2.6150 + 0.3814 = -2.2336	-1.6341 + 1.3464 = -0.2877
0	1	0	0	-2.6150 + 4.8122 = -2.1972	-1.6341 + 1.3464 = 1.9925
0	0	1	0	-2.6150 + 2.9026 = 0.2876	-1.6341 + 1.3464 = 1.8459
0	0	0	1	-2.6150 + 0 = -2.6150	-1.6341 + 0 = -1.6341

Untuk memprediksi nilai-nilai peluang, kita dapat menggunakan persamaan [1]. Misalnya, nilai prediksi peluang untuk X_1 = 1, X_2 = 0 dan X_3 = 0 adalah

$$p_1 = \frac{exp(-2.2336)}{1+exp(-2.2336)+exp(0.2877)} = 0.05769$$

$$p_2 = \frac{exp(-0.2877)}{1+exp(-2.2336)+exp(0.2877)} = 0.40384$$

Hasil komputasi ini dapat diperoleh secara mudah dengan menambahkan opsi `PROB` dalam perintah `MODEL`.

```
proc catmod data=coklat;
     direct x1-x4;
     response logits;
     model ymin = x1-x4 / noiter
freq prob;
   run;
```

Dengan menambahkan opsi `PROB`, maka SAS menampilkan output berikut sebagai tambahan dari output sebelumnya:

```
          Response Probabilities

              Response Number
 Sample  1          2          3
---------------------------------
  1    0.05769  0.15385  0.78846
  2    0.15385  0.73077  0.11538
  3    0.51923  0.42308  0.05769
  4    0.05769  0.40385  0.53846
```

Dalam contoh sebelumnya, X_1, X_2, X_3, dan X_4 adalah peubah boneka (*dummy variables*) untuk menyatakan setiap tipe aditif. Hasil yang sama juga dapat diperoleh dengan menggunakan peubah kategorik untuk menyatakan tipe aditif. Ini dikerjakan dengan cara menciptakan peubah baru X berikut:

```
   if x1=1 then x=1;
     else if x2=1 then x=2;
     else if x3=1 then x=3;
     else if x4=4 then x=4;
     else x=0;
```

Jika kode ini disisipkan dalam tahap `DATA`, kemudian kita menambahkan perintah berikut (setelah kode `proc catmod` sebelumnya):

```
proc catmod data=coklat;
     response logits;
     model ymin = x / noiter freq
prob;
      run;
```

maka SAS memberikan output:

```
                    The CATMOD Procedure

 Response            ymin       Response Levels    3
 Weight Variable     frek       Populations        4
 Data Set            COKLAT     Total Frequency  208
 Frequency Missing   0          Observations      36


                  Population Profiles
                 Sample    x    Sample Size
                 --------------------------
                     1     1             52
                     2     2             52
                     3     3             52
                     4     4             52


                     Response Profiles
                     Response    ymin
                     ---------------ƒ
                         1       1
                         2       2
                         3       3

                  Response Frequencies
                         Response Number
             Sample       1        2        3
             ---------------------------------
                 1        3       21       28
                 2       27       22        3
                 3        8       38        6
                 4        3        8       41

                 Response Probabilities
                          Response Number
          Sample         1          2          3
          ---------------------------------------
              1     0.05769    0.40385    0.53846
              2     0.51923    0.42308    0.05769
              3     0.15385    0.73077    0.11538
              4     0.05769    0.15385    0.78846

                Maximum Likelihood Analysis
         Maximum likelihood computations converged.

          Maximum Likelihood Analysis of Variance

     Source               DF   Chi-Square    Pr > ChiSq
     --------------------------------------------------
     Intercept             2        18.26        0.0001
     x                     6        78.70        <.0001
     Likelihood Ratio      0          .            .

         Analysis of Maximum Likelihood Estimates
                                      Standard      Chi-
Effect      Parameter   Estimate       Error      Square   Pr > ChiSq
-------------------------------------------------------------------
Intercept        1       -0.5909      0.2946        4.02       0.0449
                 2        0.4791      0.2242        4.57       0.0326
x                3       -1.6427      0.5209        9.95       0.0016
                 4       -0.7668      0.3032        6.39       0.0114
                 5        2.7881      0.5215       28.59       <.0001
                 6        1.5133      0.4895        9.56       0.0020
                 7        0.8786      0.4823        3.32       0.0685
                 8        1.3667      0.3831       12.73       0.0004
```

Hasil ini memperlihatkan setiap logit dugaan, yang dapat dihitung sebagai berikut:

X	$log\left(\frac{p_1}{p_3}\right)$	$log\left(\frac{p_2}{p_3}\right)$
1	–0.5909 – 1.6427 = –2.2336	0.4791 – 0.7668 = –0.2877
2	–0.5909 + 2.7881 = 2.1972	0.4791 + 1.5133 = 1.9942
3	–0.5909 + 0.8786 = 0.877	0.4791 + 1.3667 = 1.8458
4	–0.5909 – (–1.6427 + 2.7881 + 0.8786) = –2.6149	0.4791 – (– 07668 + 1.5133 + 1.3667) = –1.6341

Regresi Logit Terampat dalam SPSS: GENLOG

Dalam SPSS, model logit dikerjakan dengan prosedur GENLOG. Prosedur ini juga digunakan untuk analisis model loglinear umum (*general loglinear analysis*). Bedanya dengan SAS, SPSS 9.01 hanya dapat mengerjakan model logit multinomial yang peubah penjelasnya kategorik. Sintaks dasar model logit multinomial:

```
genlog y by x1 x2
   /model=multinomial
   /design y y*x1 y*x2 y*x1*x2.
```

dimana y adalah peubah respon, sedang x1 dan x2 adalah peubah penjelas kategorik. Data yang sama (versi yang dimodifikasi) kita gunakan untuk teladan disini:

```
genlog ymin by x
   /model=multinomial
   /print freq estim
   /plot none
   /criteria=cin(95)
iteration(20) converge(.001)
delta(0)
   /design ymin ymin*x.
```

Output dari SPSS adalah:

```
                GENERALIZED LOGLINEAR ANALYSIS

                       Data Information
                   208 cases are accepted.
        0 cases are rejected because of missing data.
      208 weighted cases will be used in the analysis.
                     12 cells are defined.
           0 structural zeros are imposed by design.
               0 sampling zeros are encountered.

                     Variable Information
                     Factor Levels Value
                          YMIN   3
                            1.00
                            2.00
                            3.00
                          X      4
                            1.00
                            2.00
                            3.00
                            4.00

                 Model and Design Information

                   Model: Multinomial Logit
                 Design: Constant + YMIN + YMIN*X

Note:  There is a separate constant term for each combination of levels of the
                      independent factors.
```

```
        Correspondence Between Parameters and Terms of the Design

                   Parameter    Aliased      Term
            1                    Constant for [X = 1.00]
            2                    Constant for [X = 2.00]
            3                    Constant for [X = 3.00]
            4                    Constant for [X = 4.00]
                 5                         [YMIN = 1.00]
                 6                         [YMIN = 2.00]
                 7         x               [YMIN = 3.00]
            8                   [YMIN = 1.00]*[X = 1.00]
            9                   [YMIN = 1.00]*[X = 2.00]
            10                  [YMIN = 1.00]*[X = 3.00]
            11       x          [YMIN = 1.00]*[X = 4.00]
            12                  [YMIN = 2.00]*[X = 1.00]
            13                  [YMIN = 2.00]*[X = 2.00]
            14                  [YMIN = 2.00]*[X = 3.00]
            15       x          [YMIN = 2.00]*[X = 4.00]
            16       x          [YMIN = 3.00]*[X = 1.00]
            17       x          [YMIN = 3.00]*[X = 2.00]
            18       x          [YMIN = 3.00]*[X = 3.00]
            19       x          [YMIN = 3.00]*[X = 4.00]

        Note: 'x' indicates an aliased (or a redundant) parameter.
                   These parameters are set to zero.

                        Convergence Information

                    Maximum number of iterations: 20
                   Relative difference tolerance: .001
                  Final relative difference: 2.92779E-14
         Maximum likelihood estimation converged at iteration 1.

                           Table Information
                        Observed        Expected
                    Factor Value    Count %   Count %
                              X      1.00
         YMIN   1.00      3.00 (  5.77)          3.00 (  5.77)
         YMIN   2.00     21.00 ( 40.38)         21.00 ( 40.38)
         YMIN   3.00     28.00 ( 53.85)         28.00 ( 53.85)
                              X      2.00
         YMIN   1.00     27.00 ( 51.92)         27.00 ( 51.92)
         YMIN   2.00     22.00 ( 42.31)         22.00 ( 42.31)
         YMIN   3.00      3.00 (  5.77)          3.00 (  5.77)
                              X      3.00
         YMIN   1.00      8.00 ( 15.38)          8.00 ( 15.38)
         YMIN   2.00     38.00 ( 73.08)         38.00 ( 73.08)
         YMIN   3.00      6.00 ( 11.54)          6.00 ( 11.54)
                              X      4.00
         YMIN   1.00      3.00 (  5.77)          3.00 (  5.77)
         YMIN   2.00      8.00 ( 15.38)          8.00 ( 15.38)
         YMIN   3.00     41.00 ( 78.85)         41.00 ( 78.85)

                        Goodness-of-fit Statistics

                     Chi-Square   DF     Sig.
              Likelihood Ratio         .0000  0      .
              Pearson                  .0000  0      .

                         Analysis of Dispersion

      Source of Dispersion      Entropy      Concentration       DF
            Due to Model         55.4019        35.2404      12
            Due to Residual     163.2376        97.3462     402
            Total               218.6395       132.5865     414

                        Measures of Association
                           Entropy = .2534
                        Concentration = .2658
```

FIBONACCI : Jurnal Pendidikan Matematika dan Matematika. Vol. 7 (2), pp: 109 – 126

```
                    Parameter Estimates
                   Constant      Estimate
                       1      3.3322
                       2      1.0986
                       3      1.7918
                       4      3.7136
Note: Constants are not parameters under multinomial assumption.
        Therefore, standard errors are not calculated.

                         Asymptotic 95% CI
  Parameter     Estimate     SE      Z-value       Lower  Upper
     5      -2.6150      .5981  -4.37         -3.79  -1.44
     6      -1.6341      .3865  -4.23         -2.39   -.88
      7         .0000      .        .            .      .
     8        .3814      .8525    .45         -1.29   2.05
     9       4.8122      .8533   5.64          3.14   6.48
     10      2.9026      .8058   3.60          1.32   4.48
      11       .0000      .        .            .      .
     12      1.3464      .4824   2.79           .40   2.29
     13      3.6266      .7268   4.99          2.20   5.05
     14      3.4800      .5851   5.95          2.33   4.63
      15       .0000      .        .            .      .
      16       .0000      .        .            .      .
      17       .0000      .        .            .      .
      18       .0000      .        .            .      .
      19       .0000      .        .            .      .
```

Output SPSS mempunyai struktur model berikut:
Anggaplah Y_{MIN}=3 sebagai kategori acuan (*reference/baseline category*), maka logit dugaan untuk Y_{MIN}=1 pada X=1 adalah

$$log\left(\frac{m_{11}}{m_{31}}\right) = \lambda^{YMIN=1} + \lambda^{YMIN=1 \text{ dan } X=1}$$

dimana m_{11} merupakan hitungan yang diprediksi untuk Y_{MIN}=1 pada X=1 dan m_{31} merupakan hitungan yang diprediksi untuk Y_{MIN}=3 pada X=1. Serupa, logit yang diduga dari Y_{MIN}=2 pada X=1 adalah

$$log\left(\frac{m_{21}}{m_{31}}\right) = \lambda^{YMIN=2} + \lambda^{YMIN=2 \text{ dan } X=1}$$

dimana m_{21} merupakan hitungan yang diprediksi untuk Y_{MIN}=2 pada X=1. Komputasi logit yang lain pada X=2, 3, dan 4 dapat diturunkan dengan cara serupa.

Output SPSS di atas menyediakan hasil pendugaan berikut:

$$log\left(\frac{m_{11}}{m_{31}}\right) = -2.6150 + 0.3814$$
$$= -2.2336, \quad log\left(\frac{m_{12}}{m_{32}}\right)$$
$$= -2.6150 + 4.8122 = 2.1972$$
$$log\left(\frac{m_{13}}{m_{33}}\right) = -2.6150 + 2.9026$$
$$= \ 0.2876, \quad log\left(\frac{m_{14}}{m_{34}}\right)$$
$$= -2.6150 + 0 = -2.6150$$

$$log\left(\frac{m_{21}}{m_{31}}\right) = -1.6341 + 1.3464$$
$$= -0.2877, \quad log\left(\frac{m_{22}}{m_{32}}\right)$$
$$= -1.6341 + 3.6266 = 1.9924$$

$$log\left(\frac{m_{23}}{m_{33}}\right) = -1.6341 + 3.4800$$
$$= \ 1.8459, \quad log\left(\frac{m_{24}}{m_{34}}\right)$$
$$= -1.6341 + 0 = -1.6341$$

Output SPSS ini tepat sama dengan yang diberikan oleh SAS.

Regresi Logit Bersyarat dalam SAS: PROC PHREG

Prosedur SAS `PHREG` mengerjakan analisis data peluang-hidup (*survival*) berdasarkan model resiko kesebandingan Cox (*Cox's proportional hazards model*). Fungsi kemungkinannya (*likelihood function*) sama dengan model logit bersyarat (*conditional logit model*). Untuk model logit bersyarat dengan `PROC PHREG`, kita perlu mengatur gugus data dalam bentuk yang sesuai untuk analisis data peluang-hidup. Dalam model ini, dianggap bahwa pilihan yang paling disukai terjadi pada waktu 1 dan semua pilihan yang lain terjadi diwaktu berikutnya atau dikatakan disensor (tidak dapat diamati). Disini perlu dibuat peubah status (*status variable*) untuk menyatakan apakah pengamatan disensor ataukah tidak. Peubah indikator mempunyai nilai 0 jika pilihan disensor dan 1 jika tidak (artinya dipilih). Sintaks dalam SAS:

```
   proc phreg;
      strata peubah_strata;
      model
peubah_waktu*peubah_status(0)  =  x1
x2;
   run;
```

dimana `peubah_strata` adalah nama dari peubah yang menentukan stratifikasi, `peubah_waktu` adalah nama dari peubah waktu gagal (nilai-nilai yang kecil berarti alternatif dipilih), `peubah_status` adalah nama dari peubah indikator yang disensor (0 jika disensor dan 1 jika tidak), sedang x1 dan x2 adalah peubah penjelas (*regressor*, *explanatory variable*, *covariate*).

Teladan berikut berasal dari Logistic Regression Examples Using the SAS System, pp. 2-3 (SAS, 1995). Sepuluh subyek (responden) disajikan 8 macam permen coklat, kemudian diminta untuk memilih satu yang paling disukai. Kedelapan permen ini merupakan kombinasi: warna (1= coklat, 0=susu), tekstur (1=lunak, 0= keras), bentuk (1=bulat, 0=tidak). Tahapan `DATA` berikut menciptakan gugus data SAS yang diberi nama `COKLAT`:

```
   data coklat;
      input subyek pilihan warna
tekstur bentuk @@;
      t=2-pilihan;
      datalines;
   1 0 0 0 0   1 0 0 0 1   1 0 0 1 0
1 0 0 1 1
    1 1 1 0 0   1 0 1 0 1   1 0 1 1 0
1 0 1 1 1
    2 0 0 0 0   2 0 0 0 1   2 0 0 1 0
2 0 0 1 1
    2 0 1 0 0   2 1 1 0 1   2 0 1 1 0
2 0 1 1 1
    3 0 0 0 0   3 0 0 0 1   3 0 0 1 0
3 0 0 1 1
    3 0 1 0 0   3 0 1 0 1   3 1 1 1 0
3 0 1 1 1
    4 0 0 0 0   4 0 0 0 1   4 0 0 1 0
4 0 0 1 1
    4 1 1 0 0   4 0 1 0 1   4 0 1 1 0
4 0 1 1 1
    5 0 0 0 0   5 1 0 0 1   5 0 0 1 0
5 0 0 1 1
    5 0 1 0 0   5 0 1 0 1   5 0 1 1 0
5 0 1 1 1
    6 0 0 0 0   6 0 0 0 1   6 0 0 1 0
6 0 0 1 1
    6 0 1 0 0   6 1 1 0 1   6 0 1 1 0
6 0 1 1 1
    7 0 0 0 0   7 1 0 0 1   7 0 0 1 0
7 0 0 1 1
    7 0 1 0 0   7 0 1 0 1   7 0 1 1 0
7 0 1 1 1
    8 0 0 0 0   8 0 0 0 1   8 0 0 1 0
8 0 0 1 1
    8 0 1 0 0   8 1 1 0 1   8 0 1 1 0
8 0 1 1 1
    9 0 0 0 0   9 0 0 0 1   9 0 0 1 0
9 0 0 1 1
    9 0 1 0 0   9 1 1 0 1   9 0 1 1 0
9 0 1 1 1
   10 0 0 0 0  10 0 0 0 1  10 0 0 1 0
10 0 0 1 1
   10 0 1 0 0  10 1 1 0 1  10 0 1 1 0
10 0 1 1 1
   ;
```

dimana subyek sebagai nomor subyek (*subject number*), `pilihan` sebagai peubah status (*status variable*), dan t sebagai peubah waktu (*time variable*). Kita menggunakan sintaks berikut:

```
   proc phreg data=coklat;
      strata subyek;
      model t*pilihan(0)=warna
tekstur bentuk;
   run;
```

Hasil analisis adalah sebagai berikut:

```
Analisis Data Peluang-hidup: The PHREG Procedure

Model Information

Data Set                  WORK.COKLAT
    Dependent Variable       t
  Censoring Variable       pilihan
    Censoring Value(s)       0
  Ties Handling            BRESLOW

Summary of the Number of Event and Censored Values
                                  Percent
Stratum    subyek     Total     Event    Censored    Censored
   1          1          8         1          7        87.50
   2          2          8         1          7        87.50
   3          3          8         1          7        87.50
   4          4          8         1          7        87.50
   5          5          8         1          7        87.50
   6          6          8         1          7        87.50
   7          7          8         1          7        87.50
   8          8          8         1          7        87.50
   9          9          8         1          7        87.50
  10         10          8         1          7        87.50
 Total                  80        10         70        87.50

Convergence Status: Convergence criterion (GCONV=1E-8) satisfied.


Model Fit Statistics
         Without          With
Criterion     Covariates     Covariates
-2 LOG L          41.589         28.727
AIC               41.589         34.727
SBC               41.589         35.635


Testing Global Null Hypothesis: BETA=0

Test                  Chi-Square       DF     Pr > ChiSq
Likelihood Ratio         12.8618        3         0.0049
Score                    11.6000        3         0.0089
Wald                      8.9275        3         0.0303


Analysis of Maximum Likelihood Estimates

         Parameter    Standard                                   Hazard
Variable  DF    Estimate       Error   Chi-Square   Pr > ChiSq     Ratio
warna      1     1.38629     0.79057       3.0749       0.0795     4.000
tekstur    1    -2.19722     1.05409       4.3450       0.0371     0.111
bentuk     1     0.84730     0.69007       1.5076       0.2195     2.333
```

Hasil di atas memberikan model pendugaan peluang berikut:

$$p_j = \frac{exp(1.38629 * \text{warna}_j - 2.19722 * \text{tekstur}_j + 0.84730 * \text{bentuk}_j)}{\sum_{j=1}^{n} exp(1.38629 * \text{warna}_j - 2.19722 * \text{tekstur}_j + 0.84730 * \text{bentuk}_j)}, \text{ dimana } n = 8.$$

Nilai dugaan yang positif untuk warna dan bentuk berarti kedua peubah ini meningkatkan preferensi, sedang nilai dugaan negatif (tekstur) menurunkan preferensi. Jadi, macam permen yang paling disukai oleh konsumen adalah yang punya warna coklat, tekstur keras, dan bentuk bulat. Ini diperlihatkan oleh nilai prediksi peluang berikut:

Pilih-an	Warna	Tekstur	Bentuk	$exp(x_j^T\beta_j)$	Pred. Peluang
1	0	0	0	1.000	0.054
2	0	0	1	2.333	0.126
3	0	1	0	0.111	0.006
4	0	1	1	0.269	0.014
5	1	0	0	4.000	0.216
6	1	0	1	**9.333**	**0.504**
7	1	1	0	0.444	0.024
8	1	1	1	1.037	0.056
				18.518	1.000

Regresi Logit Bersyarat dalam SPSS: Prosedur COXREG

Logit bersyarat dalam SPSS dikerjakan dengan menggunakan prosedur COXREG yang sintaksnya sebagai berikut:

```
coxreg peubah_waktu with X1 X2
  /status=peubah_status(1)
  /strata=peubah_strata.
```

dimana `peubah_waktu` adalah nama dari peubah waktu gagal (nilai yang lebih kecil berarti alternatif dipilih), `peubah_status` adalah nama dari peubah indikator yang disensor (1=disensor, 0=tidak), `peubah_strata` adalah nama dari peubah yang menentukan stratifikasi, sedang x_1 dan x_2 adalah peubah penjelas.

Untuk data kita, kode program SPSS adalah

```
coxreg t with warna tektur bentuk
     /status=pilihan(1)
     /strata=subyek.
```

yang menghasilkan output sebagai berikut:

```
                    C O X R E G R E S S I O N

                      80 Total cases read
                   0 Cases with missing values
               0 Valid cases with non-positive times
        0 Censored cases before the earliest event in a stratum
                      0 Total cases dropped
                80 Cases available for the analysis

                     Dependent Variable: T

                 SUBJECT      Events  Censored
               1.00          1         7 (87.5%)
               2.00          1         7 (87.5%)
               3.00          1         7 (87.5%)
               4.00          1         7 (87.5%)
               5.00          1         7 (87.5%)
               6.00          1         7 (87.5%)
               7.00          1         7 (87.5%)
               8.00          1         7 (87.5%)
               9.00          1         7 (87.5%)
               10.00         1         7 (87.5%)
                Total          10     70 (87.5%)

      Beginning Block Number 0. Initial Log Likelihood Function

                 -2 Log Likelihood   41.589
```

```
                  Beginning Block Number 1.  Method: Enter

                   Variable(s) Entered at Step Number 1..
                                WARNA
                                BENTUK
                                TEKSTUR

                 Coefficients converged after 5 iterations.

                        1-2 Log Likelihood  28.727


                       Chi-Square   df     Sig
         Overall (score)     11.600            3      .0089
                         Change (-2LL) from
         Previous Block      12.862            3      .0049
         Previous Step       12.862            3      .0049

------------------------ Variables in the Equation ---------------------
 Variable        B       S.E.    Wald     df   Sig     R        Exp(B)
 WARNA        1.3863   .7906   3.0749    1   .0795   .1608    4.0000
 BENTUK        .8473   .6901   1.5076    1   .2195   .0000    2.3333
 TEKSTUR     -2.1972  1.0541   4.3450    1   .0371   -.2375    .1111

                          Covariate Means
                         Variable      Mean
                         WARNA         .5000
                         BENTUK        .5000
                         TEKSTUR       .5000
```

Hasil pendugaan ini tepat sama dengan hasil PROC PHREG SAS.

SIMPULAN

Dua model regresi logit multinomial untuk peubah respon politomus atau multinomial yang strukturnya tak berurut, yaitu (i) model logit terampat (*generalized*) dan (ii) model logit bersyarat. Model logit menggunakan karakteristik dari individu (subyek) sebagai peubah penjelas, sedang model logit bersyarat menggunakan karakteristik dari pilihan individu. Dalam dunia pendidikan karakteristik pilihan individu ini dapat berwujud penilaian pada rubrik penilaian yang berupa pilihan preferensi psikologis atau angka tingkat kesetujuan/ketidaksetujuan terhadap suatu hal yang dievaluasi. Sedangkan karakter individu pada model regresi logit terampat dapat berupa latar belakang pendidikan dan sosio-kultur-ekonomi dari penilai dalam hal ini bisa guru, dosen atau asesor.

Regresi logit terampat (*generalized logit*) dapat dilakukan dengan PROC CATMOD dalam SAS. Sedangkan untuk melakukan perhitungan model regresi logit bersyarat (*conditional logit*) terdapat dua pilihan yaitu dengan SAS melalui PROC PHREG atau dengan SPSS melalui PROC COXREG. Kedua pendekatan untuk regresi logit bersyarat ini telah ditunjukkan dalam artikel ini menghasilkan luaran yang sama.

Lebih jauh, meskipun di luar cakupan pembahasan kali ini, penelaahan terhadap penggabungan kedua model logit ini menurut hemat penulis sangat menarik untuk dilakukan. Yaitu sebuah upaya untuk melakukan pemodelan yang menggunakan karakteristik dari individu (subyek) sebagai peubah penjelas seperti pada model logit terampat, dan sekaligus menggunakan karakteristik dari pilihan individu seperti pada model logit bersyarat.

DAFTAR PUSTAKA

Agresti, Alan. 1996. *An Introduction to Catagorical Data Analysis*. New York: John Wiley & Sons, Inc.

Agresti, Alan. 1990. *Catagorical Data Analysis*. New York: John Wiley & Sons, Inc.

Aitkin, M., Anderson, D., Francis, B., & Hinde, J. 1989. *Statistical Modelling in GLIM*. New York: Clarenden Press.

Hadi, A. F. 2021. Statistika Intuitif. Elstat: Statistika Digital Indonesia. [Online] Tersedia: https://elstat.id/intutivestatistics-an-introduction-1/ [1 Desember 2021].

Guido, J. J, Winters, P. C., & Rains, A. B. 2006. *Logistic Regression Basics*. NESUG 2006. [Online] Tersedia: https://www.lexjansen.com/nesug/nesug06/an/da26.pdf [10 November 2021].

SAS Institute. 2000. SAS/Stat Software: Changes and Enhancements through Release 6.11 Paperback – August 1, 2000.

SAS Technical Support. 2005. [Online] Tersedia: http://support.sas.com/techsup/tnote/tnote_stat.html [10 November 2021]

McCullagh, P., & Nelder, J. A. 1989. *Generalized Linear Models*. London: Chapman & Hall.

Ying So, and Kuhfeld, W. F. 1995. *Multinomial logit models*, Makalah disajikan dalam SAS Users Group International 20 (SUGI 20). Orlando, Florida, 2-5 April.