Jurnal Penelitian Sains; hal 1 - 6
No.1. Oktober 1996 ISSN:

KAJIAN PENGKODEAN VARIABEL KUALITATIF NOMINAL

A. Ramali L.P., E. S. Cahyono, dan B. Suprihatin
Jurusan Matematika FMIPA Universitas Sriwijaya

ABSTRAK

Penelitian ini bertujuan untuk mengkaji rincian pembentukan pengkodean dan perumusan teorema variabel kualitatif nominal multipel. Selain itu, melakukan pembuktian teorema yang dirumuskan untuk variabel kualitatif nominal yang diperoleh. Hasil akhir yang diperoleh adalah rincian bentuk umum dari teorema pengkodean dengan pembuktiannya serta teorema dan definisi yang mendukung pembuktian.

PENDAHULUAN

Kuantifikasi data variabel kualitatif merupakan langkah pertama yang dilakukan untuk data variabel kualitatif, agar data tersebut dapat dianalisis secara statistika, sehingga hasil analisisnya dapat memberikan banyak informasi dalam bentuk parameter-parameter yang dapat menjelaskan berbagai aspek yang dimiliki oleh keseluruhan unit penelitian.

Sejauh ini, kuantifikasi data variabel kualitatif, khususnya nominal, menggunakan pengkodean (*coding*) numerik dari *modalitas* variabel kualitatif nominal (Caillez dan Pages, 1976 ; hal. 191,401). Teorema mengenai pengkodean dirumuskan secara eksplisit untuk suatu himpunan modalitas.

Pada penelitian ini, akan dicoba untuk mengkaji teorema pengkodean variabel kualitatif nominal yang berlaku untuk satu atau lebih himpunan modalitas, yang perumusannya secara umum dapat dilihat pada Djauhari (1988), sehingga akan dapat diketahui rincian pembuktian ataupun teorema dan definisi pendukungnya.

Hal ini perlu dilakukan, karena pada kenyataannya masalah yang akan dianalisis secara statistika hampir selalu melibatkan sejumlah variabel acak pengamatan sekaligus. Dengan demikian, kajian ini dapat menjadi pijakan untuk memahami proses pengkualifikasian data melalui pengkodean variabel kualitatif nominal yang dapandang dari sudut matematiknya (aljabar linier).

HASIL DAN PEMBAHASAN

Kuantifikasi Variabel Kualitatif Nominal Multipel

Kuantifikasi variabel kualitatif nominal untuk lebih dari satu variabel dapat dirumuskan dan dirinci sebagai berikut: Misalkan populasi E yang diamati memiliki nilai Card (E) = p_j, dengan uraian sebagai berikut:

$$
\begin{array}{llllll}
X_1 & x_{11}, & x_{12}, & \ldots, & x_{1p_1} & \text{dengan modalitas } Q_1 = \{1,2,\ldots,p_1\}. \\
X_2 & x_{21}, & x_{22}, & \ldots, & x_{2p_2} & \text{dengan modalitas } Q_2 = \{1,2,\ldots,p_2\}. \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
X_j & x_{j1}, & x_{j2}, & \ldots, & x_{jp_j} & \text{dengan modalitas } Q_j = \{1,2,\ldots,p_j\}. \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
X_m & x_{m1}, & x_{m2}, & \ldots, & x_{mp_m} & \text{dengan modalitas } Q_m = \{1,2,\ldots,p_m\}.
\end{array}
$$

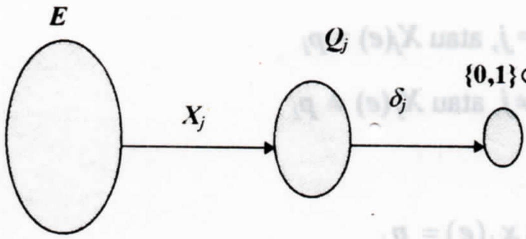
adalah variabel-variabel indikator dari X_j, kemudian nyatakan X_j sebagai variabel acak kualitatif nominal dengan Q_j menyatakan modalitas ke-j, maka variabel statistik X_j dapat dipandang sebagai pemetaan $X : E \longrightarrow Q_j$. Himpunan Q_j bukan merupakan himpunan bagian dari himpunan bilangan riil $\Re$, atau $Q_j \not\subset \Re$. Selanjutnya, jika ingin melakukan kuantifikasi X, maka yang diperlukan adalah sebuah pemetaan $\delta_j : Q_j \longrightarrow \Re$, sehingga $\delta_1 . X_1, \delta_2 . X_2, \ldots, \delta_j . X_j$ dengan δ_j sebagai *coding* untuk X_j. Jadi $\delta_j . X_j$ merupakan bentuk kuantifikasi dari X_j.

Dengan demikian, proses kuantifikasi variabel acak kualitatif nominal di atas dapat digambarkan seperti di bawah ini.

$$\begin{tikzcd} E \arrow[r, "X_j"] \arrow[dr, "\delta_j \,.\, X_j"'] & Q_j \arrow[d, "\delta_j"] \\ & \{0,1\} \subset \Re \end{tikzcd}$$

Gambar 1. Pemetaan Variabel Kualitatif Nominal
untuk Variabel Multipel.

Atau, dapat juga dibuat bentuk lain dari pemetaan di atas seperti di bawah ini



Gambar 2. Pemetaan Variabel Kualitatif Nominal
untuk Variabel Multipel.

Selanjutnya, δ_j disebut *kode* (*coding*) dari variabel acak kualitatif nominal X_j, untuk mengetahui apa yang dimaksud dengan $\delta_j \,.\, X_j$, digunakan teorema berikut ini.

Pembentukan Teorema Pengkodean bagi Variabel Kualitatif Nominal Multipel

Pembentukan dan rincian pembuktian teorema pengkodean bagi variabel kualitatif nominal multipel merupakan rumusan umum dari kuantifikasi variabel kualitatif nominal multipel di atas yang dilakukan seperti berikut dibawah ini.

Teorema 1. Misalkan variabel acak kualitatif X_j memiliki himpunan modalitas $Q_j = (1, 2, \ldots, P_j)$ dengan $x_{j1}, x_{j2}, \ldots, x_{jk}, \ldots x_{jp_j}$, adalah indikator dari X_j dan δ_j *coding* untuk X_j yang berlaku untuk $j = 1, 2, \ldots, m$, $i = 1, 2, \ldots, p$, dan $k = 1, 2, \ldots, p_j$ artinya untuk semua e di E, berlaku

$$x_{jk}(e) = \begin{cases} 1 \text{ ; Jika } x_j(e) = p_j \\ 0 \text{ ; Jika } x_j(e) \neq p_j \end{cases}$$

jika $\delta_j : Q_j \longrightarrow \Re$ merupakan suatu pemetaan, maka

$$\delta_j . X_j = \sum_{k=1}^{p_j} \delta_j(k) x_{jk}$$

Bukti : Ambil sembarang e di E dan misalkan $X(e) = i$, untuk suatu i di Q, dengan $i = 1,2,\ldots,p$. Misalkan $X(i) = p_j$, untuk suatu p_j di Q_j dengan $j = 1,2,\ldots,m$, $i = 1,2,\ldots,p$, dan $k = 1,2,\ldots,p_j$. Jadi $\delta_j . X_j(e)$, melalui *sifat fungsi komposit* diperoleh $\delta_j . X_j(e) = \delta_j \{X_j(e)\}$, atau $\delta_j \{X_j(e)\} = \delta_j (p_j)$.

jadi

$x_{jk}(e) = 1$; jika $k = j$, atau $X_j(e) = p_j$

$x_{jk}(e) = 0$; jika $k \neq j$, atau $X_j(e) \neq p_j$

atau dapat ditulis

$$x_{jk}(e) = \begin{cases} 1 \text{ ; jika } x_j(e) = p_j \\ 0 \text{ ; jika } x_j(e) \neq p_j \end{cases}$$

akibatnya $\delta_j . X_j(e) = \delta_j . X_j(e) = \delta_j \{X_j(e)\}$, atau $\delta_j \{X_j(e)\} = \delta(p_j)$. Oleh karena itu, berdasarkan *sifat perkalian vektor* (Tucker, 1989; hal 72), maka $\delta_j . X_j(e) = \delta_j . X_j(e) = \delta_j\{X(e)\} = \delta_j(p_j)$ akan sama dengan

$$\delta_j . X_j = \sum_{k=1}^{p_j} \delta_j(k) x_{jk}$$

Teorema 2. Misalkan X_j variabel kualitatif nominal dengan p_j modalitas, jika

$$\sum_{i=1}^{p_j} \delta_j(i) = 0, \text{ maka } \delta_j . X_j = \sum_{k=1}^{p_j} \delta_j(k) x_{jh} \quad \text{, atau}$$

$$\delta_j . X_j = \sum_{k=1}^{p_j - 1} \delta(k)(x_{jk} - x_{jp_j})$$

Bukti: Jika X_j variabel kualitatif nominal dengan p_j modalitas, maka δ_j akan terpusat (*centered*), artinya $\sum_{k=1}^{p_j} \delta_j(k) = 0$, dimana $\delta_j . X_j$ merupakan kombinasi linier dari variabel indikator, dengan defenisi variabel indikator adalah $X_i : X_i = 1$ jika $X = i$ dan $X_i = 0$ jika $X \neq i$ (Bouroche, 1977; hal. 136). Jadi jelas bahwa jika diketahui

$$\sum_{k=1}^{p_j} \delta_j(k) = 0.$$

maka

$$\delta_j(p_j) = -\sum_{k=1}^{p_j-1} \delta_j(k)$$

$$\delta_j(p_j) x_{jp_j} = -\sum_{k=1}^{p_j-1} \delta_j(k) x_{jp_j}.$$

$$\sum_{k=1}^{p_j-1} \delta_j(k) x_{jk} + \delta_j(p_j) x_{jk} = \sum_{k=1}^{p_j-1} \delta_j(k) x_{jk} - \sum_{k=1}^{p_j-1} \delta_j(k) x_{jp_j}$$

Jadi

$$\sum_{k=1}^{p_j} \delta_j(k) x_{jk} = \sum_{j=1}^{m} \sum_{k=1}^{p_j-1} \delta_j(k)(x_{jk} - x_{jp_j}).$$

$$\delta_j . X_j = \sum_{k=1}^{p_j} \delta_j(k) x_{jk} = \sum_{j=1}^{m} \sum_{k=1}^{p_j-1} \delta_j(k)(x_{jk} - x_{jp_j}),$$

atau

$$\delta_j . X_j = \sum_{k=1}^{p_j-1} \delta(k)(x_{jk} - x_{jp_j}),$$

untuk $j = 1,2,...,m$, $i = 1,2,...,p$-1 dan $k = 1,2,..., p_j$-1

KESIMPULAN

Hasil yang diperoleh dari penelitian ini adalah rincian bentuk umum teorema pengkodean variabel kualitatif nominal serta sejumlah defenisi dan teorema pendukung yang digunakan dalam menguraikan dan membuktikan teorema pengkodean yang diperoleh. Selanjutnya disarankan untuk mengkaji lebih jauh syarat harga δ jika tidak terpusat (*non centered*). Selain itu, dapat diperluas dengan membangun teorema pengkodean variabel kualitatif ordinal.

DAFTAR PUSTAKA

Bourouche, T.M. 1977. Analyse des Donnees en Marketing. Masson. Paris.

Caillez, F. dan Pages, J.P. 1976. Introduction a L'analyse des Donnees. Smash. Paris.

Djauhari, M.A. 1988. Struktur Data Statistik. Penerbit Karunika. Jakarta.

Tucker, A. 1989. A Unified Introduction to Linear Algebra Model. Model Methods, and Theory. Mac Millan Publishing Company. New York.