II. TINJAUAN PUSTAKA

2.1 Proses Stokastik

Stokastik proses $\underline{X} = \{X(t), t \in T\}$ adalah kumpulan dari variabel acak yang didefinisikan pada ruang peluang (Ω, ς, P) yang nilai-nilainya pada bilangan real. T dinamakan himpunan indeks dari proses atau ruang parameter yang biasanya adalah subset dari R. Himpunan nilai-nilai dimana variabel acak $X(t)$ dinamakan *state space* (ruang keadaan) dari proses dan biasanya dilambangkan dengan S. Pemetaan untuk setiap $\omega \in \Omega$:

$$X_{(\omega)}: \; T \to S \; ,$$

$$t \; \to X_t(\omega)$$

Dinamakan *simple path* (Castadena, *et al*., 2012).

2.2 Rantai Markov Waktu Diskrit

Rangkaian dari variabel acak $(X_n)_{n \in N}$ dengan ruang keadaan diskrit dikatakan rantai markov waktu diskrit jika memenuhi :

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P\,(X_{n+1} = j \mid X_n = i)$$

Untuk semua $n \in N$ dan untuk semua $i_0, i_1, \ldots, i_{n-1}, i, j \in S$ dengan :

$$P\ (X_0 = i_0, \ldots, X_n = i_n) > 0$$

Dengan kata lain, kondisi tersebut menyatakan secara tidak langsung jika *state* sekarang diketahui "$X_n = i_n$", maka *state* sebelumnya "$X_{n-1}, X_{n-2}, \ldots, X_0$" tidak mempengaruhi peluang *state* yang akan datang X_{n+1} (Castadena, *et al.*, 2012).

2.3 Matriks Peluang Transisi

Castadena, *et al*. (2012) mendefinisikan matriks peluang transisi sebagai :

$$P = (p_{ij}) = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Dengan catatan :

$p_{ij} \geq 0$ *untuk semua* $i, j \in S$ dan $\sum_j p_{ij} = 1$ *untuk semua* $i \in S$.

2.4 Analisis Rantai Markov

Penentuan karakteristik dari suatu proses stokastik dapat dilakukan melalui klasifikasi rantai markov. Pada penelitian ini akan ditentukan peluang stasioner yang akan mewakili karakteristik data migrasi Indonesia. Sehingga dihasilkan peluang konstan untuk setiap waktunya. Berikut adalah beberapa klasifikasi rantai markov menurut Sheldon (1996).

Definisi 2.1

State j dikatakan dapat dicapai (*accessible*) dari *state* i dengan n ≥ 0 steps jika $P_{ij}^{(n)} > 0$. Dapat ditulis dengan $i \rightarrow j$ (j dapat dicapai dari i dengan n langkah).

Definisi 2.2

State i dan j berkomunikasi jika $i \rightarrow j$ dan $j \rightarrow i$. Dapat disimbolkan dengan $i \leftrightarrow j$.

Definisi 2.3

Rantai markov dikatakan *irreducible* jika ruang *state* terdiri dari hanya 1 kelas, artinya semua *state* berkomunikasi satu sama lain. Dapat dengan mudah bahwa $i \leftrightarrow j$ merupakan relasi ekivalen pada S sehingga kelas ekivalen $C(i) := \{j \in S : i \leftrightarrow j\}, i \in S$, dari partisi S.

Definisi 2.4

Misalkan $i \in S$ tetap. Perioda i didefinisikan sebagai berikut:

$$\lambda_i = GCD\ \{n \geq 1 : p_{ii}{}^{n} > 0\} \qquad (2.1)$$

Dimana GCD adalah faktor persekutuan terbesar (FPB). Jika $p_{ii}{}^{n} = 0$ untuk semua n ≥ 1, maka $\lambda_i = 0$.

Definisi 2.5

State i dikatakan aperiodik ketika $\lambda_i = 1$.

Definisi 2.6

State j dikatakan *persistent* atau *recurrent* jika $f_{jj} = 1$ (yaitu pasti kembali ke *state* j). Selain itu, semua *state* yang *recurrent* pada rantai markov dengan *state space* berhingga adalah *recurrent* positif.

Definisi 2.7

Ambil $\{X_n; n \geq 0\}$ adalah rantai markov dengan ruang *state* S. Peluang $\pi=(\pi_h)_{h\epsilon S}$ dikatakan invarian atau stasioner jika:

$$\pi_j = \sum_{i\epsilon S} \pi_i p_{ij} \tag{2.2}$$

untuk semua $j \epsilon S$. dengan kata lain, π adalah invarian jika $\pi = \pi P$.

Teorema 2.1

Ambil $\{X_n; n \geq 0\}$ bersifat *irreducible* (semua *state* berkomunikasi satu sama lain), rantai markov aperiodik dengan ruang *state* S. Maka terdapat peluang invarian terhadap S jika hanya jika rantainya merupakan rantai *reccurent* positif.

2.5 Analisis Chi-Square

Asumsikan bahwa masing-masing individu pada setiap I populasi berada tepat pada satu kategori J. Misalkan sampel dari n_i individu yang diambil dari populasi ke-i, dengan:

$n = \sum n_i$

N_{ij} = jumlah individu sampel ke-i yang terdapat pada kolom *j*,

$N_{.j} = \sum_{i=1}^{I} N_{ij}$ = jumlah total individu antara *n* sampel yang masuk dalam kategori *j*.

N_{ij} yang teramati tercatat pada tabel kontingensi dua arah dengan baris I dan kolom J. Jumlah pada baris ke-i adalah n_i , sementara jumlah dari entri pada kolom ke-j adalah $N_{.j}$. Misalkan p_{ij} merupakan proporsi individu di populasi *i* yang masuk dalam kategori, dengan demikian untuk populasi pertama proporsi *j* adalah $p_{11}, p_{12}, \ldots, p_{1j}$ (dimana jumlahnya adalah satu) dan berlaku untuk populasi lainnya.

Hipotesis nol menyatakan bahwa proporsi individu dalam kategori J adalah sama untuk setiap populasi, dan bahwa ini adalah benar untuk setiap populasinya. Berikut adalah hipotesis untuk setiap kategori :

$H_0 : p_{1j} = p_{2j} = \cdots = p_{ij} \qquad j = 1,2, \ldots, J$

H_1 : Tidak semua proporsi setiap populasi sama

Ketika H_0 benar, maka $p_1, p_2, \ldots, p_j$ dapat digunakan dalam menotasikan proporsi populasi pada setiap J kategori. Proporsi ini umum untuk semua populasi I. Jumlah yang diharapkan dari individu pada sampel ke-i yang terdapat pada J kategori ketika H_0 benar adalah $E(N_{ij}) = n_{i}.p_{j}$. Sebelum memperkirakan $E(N_{ij})$ terlebih dahulu harus memperkirakan p_j , proporsi pada kategori *j*. Berdasarkan *n* total individu dan $N_{.j}$ masuk dalam kategori j, maka $\widehat{p_j} = \frac{N_{.j}}{n}$ sebagai penduga.

Sehingga :

$$\widehat{E_{ij}} = n_{i}.\frac{N_{.j}}{n} = \frac{(jumlah\ baris\ ke - i)(\ jumlah\ kolom\ ke - j)}{n}$$

Ketika H_0 benar , maka uji statistik:

$$X^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(N_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

berdistribusi *Chi-Square* dengan (I-1) (J-1) derajat bebas. Uji untuk H_0 dengan tingkat kepercayaan α untuk tolak H_0 jika $X^2 \geq X^2{}_{\alpha,(I-1)(J-1)}$ (Devore,1982).

Penentuan tingkat perbedaan antarproporsi dalam analisis *Chi-Square* dapat diukur dengan menggunakan nilai *Cramer's V*. Rumus *Cramer's V* yaitu:

$$V = \sqrt{\frac{X^2}{N\min(I-1, J-1)}}$$

Dimana I merupakan jumlah baris , J adalah jumlah kolom, dan N merupakan jumlah kejadian. Dalam penelitian ini, *Cramer's V* digunakan dalam mengukur tingkat perbedaan antarproporsi, semakin kecil nilai *Cramer's V* berarti semakin rendah tingkat perbedaan antarproporsi, tetapi apabila nilai *Cramer's V* mendekati satu maka semakin tinggi tingkat perbedaannya.