응용통계연구 제 7권 2호
1994년 9월 pp. 249-261

한국의 인구곡선 추정에 관한 연구[1]

구자흥[2]

요약

본 연구의 목적은 한국인구의 연도별추정과 그 장래인구의 예측을 위한 단순 Logistic Curve의 모수(母數)들을 추정하는데 있다. 즉 1949년 부터 1990년에 이르는 9회에 걸친 인구 Census자료를 토대로하고, 3-군법(3-群法)을 적용하여 모수들을 추정하므로서 연도별 총인구들의 곡선추계를 위한 Logistic Curve를 얻었다. 아울러 이 곡선에 의하여 1950년부터 1990년까지의 Census와 Census 사이의 총인구를 연도별로 추정하였으며, 다른한편 1991년부터 2010년에 이르기까지 장래인구를 예측하였다. 뿐만아니라 최근 한국인구 성장의 상한점근치(Upper Asymptote)로 58,616천명을 제시할 수 있었다.

1. 서론

인구추정 및 장래예측을 위한 이론적 방법으로서는 다음 3가지를 들수있다.

(a) 수리적 추계방법 (Mathematical Methods)
(b) 요인별 추계방법 (Component Projection)
(c) 경제적 추계방법 (Economic Methods)

그 첫번째인 수리적 방법은 예컨대, 인구Census자료와 같은 과거의 실측치인 인구의 시계열적자료에 적당한 수리모형(Mathematical Model)을 맞추어 그것에 의하여 인구의 추정(계)과 그 장래 인구 예측치들을 구하는 방법론이다.

이때 수리모형들로서는

① 선형보간법(Linear Interpolation Method)
② 2차 다항식법(Quadratic Polynomial Method)
③ 복리법(Compound Interest Method)
④ 로지스틱 곡선법(Logistic Curve Method)

등을 들수 있다.

그리고 인구추계 및 그 장래인구 예측의 대상으로서는 인구Census와 그 다음 인구Census 사이의 연도별 총 인구와 지역인구, 남・여별 인구가 그 추정 대상이 되고 장래인구 예측의 대상은 최근의 인구Census 이후 연도별 인구예측이 그 대상이 된다. 특히, 인구추계치와 예측치들은 여러가지 정책결정의 기본자료로 쓰이게 되며 또한 장래추계 인구는 경제개발계획의

---

1) 본 연구는 1992년도 인하대학교 교내연구비에 의하여 이루어 진것임.
2) (402-751) 인천직할시 남구 용현동 253, 인하대학교 통계학과.

수립에 기초자료가 되고, 미래에 닥쳐올 인구문제에 대처하거나 해소하기위한 청사진(인구정책 및 인구대책)의 수립에 필수적인 기본 정보 자료가 된다.

위 수리모형들 중에서 인구Census 구간사이의 인구(Intercensal Population)을 추정하는데는 비교적 짧은기간이므로 ①,②,③등의 방법이 흔히 이용되고, 장기예측 수리모형으로서는 ④의 Logistic곡선법이 빈번히 이용된다.

본 연구에서는 한국인구의 인구Census구간 사이의 연도별 인구추계와 인구의 장기예측 (Population Projection)을 위하여 우리나라 인구성장 곡선으로서 Logistic Curve를 그 수리모형 (Mathematical Model)으로 하고 그 곡선의 모수(Parameter) N, k, λ등을 추정하므로서 추정곡선 :

$$\widehat{P_t} = \frac{\widehat{N}}{1 + e^{\ln \widehat{k} - \widehat{\lambda} t}} \tag{1}$$

를 결정하여 보았다. 여기서 $\widehat{N}$, $\widehat{k}$, 그리고 $\widehat{\lambda}$등은 각각 모수 N, k, 그리고 λ등의 추정치를 나타낸다.

또 (1)식에 의하여 1995년, 2000년 및 2010년 등의 인구를 예측(Projection)하여 보았으며 인구증가의 상한점근치(Upper Asymptote)인 종국인구 (Ultimate Population)를 구하여 제시하였다.[3]

특히 우리나라 인구의 추정을 위하여 Logistic Curve를 채택하고, 그 모수추정을 위하여 3군법을 이용할경우 Census 자료가 풍부하여야하고, 또 그들 자료가 시계열적으로 등간격(等間隔)이어야하는데 우리의 경우 현대적 수단(통계적조직과 자료처리조직)을 동원한 인구Census의 역사가 짧고, 더욱이 1950년 한국전쟁으로인하여 인구Census가 결여되어 있으며, 1965년에 실시되어야할 인구Census가 정치 경제적여건으로 1966년에 실시된 점 등을 교정 보완이 불가피 하였다. 따라서, 1950년의 총인구 추정치를 1949년 과 1955년의 Census인구들을 이용하여 선형보간법으로 구하여 이용하였으며, 1965년도 총인구의 자료로서는 앙년추계인구(Mid-Year Estimated Population)을 이용하였다.

그밖에도 인구Census 실사(實査) 기준일자들도 9월 1일, 10월 1일, 11월 1일, 12월 1일등 다양했지만 편의상 이것들을 동일시 하였다.

2. Logistic곡선과 그 모수 추정

(a) Logistic곡선의 이론적 배경

일찌기 Pritchett는 1790년부터 1880년까지의 미국 인구Census자료들을 3차곡선에 fitting시키고, 이 곡선에 의거하여 그후 장래인구를 예측(Projection)하여 제시 한 바 있다.

다른한편, 미세한 시간 (Δt)사이의 인구증분 (ΔP)는 총인구 P의 함수로 표현 할 수 있다고 가정한다. 그러면, 다음 미분방정식을 인구성장법칙의 수리적표현으로 삼을 수 있다.

$$\frac{dP}{dt} = \psi(P) \qquad (2)$$

좀 더 현실성있게 (2)식의 비동차항을

$$\psi(P) = \mu P \qquad (3)$$

라고 두면 선형미분방정식:

$$\frac{dP}{dt} = \mu P \qquad (4)$$

를 얻는다. 그러므로 얻어진 미분방정식 (4)식을 풀어 그 일반해를 구하면 인구에 관한 지수곡선군(exponential curve family):

$$P_t = c\,e^{\mu t} \qquad (5)$$

를 얻는다.

또 초기조건(Initial Condition)을 $t=0$으로 두면 초기인구(Initial Population) $P_o=c$를 얻는다. 즉

$$P_t = P_o e^{\mu t} \qquad (6)$$

분명히 (5)식 또는 (6)식에 의하면 인구 크기는 기하급수적 성장을 하게 됨을 알수있다. 이러한 기하급수적 인구성장법칙은 말사스의 성장법칙(Malthusian Law of Growth)에서 연유 된 것이다.[3]

그러나 기하급수적 인구성장법칙에 의한 인구예측(Population Projection)은 대개의 경우 거의 불가능하다. 그 이유는 양의 증가율 $\mu(>0)$에 의한 장기간의 경우 엄청난 인구크기(population size)을 보여주기 때문이다. 그러므로 인구의 장기예측을 위한 수리적표현으로서 가장 많이 쓰이는 성장곡선으로 Logistic곡선을 들수있다.

1838년 벨기엘의 수학자 P. F. Verhulst는 (3)식의 비동차항을 다음과 같이 두고 Logistic곡선을 창출하였다.

$$\psi(P) = \mu P - \phi(P) \qquad (7)$$

그리고 또 다시

$$\phi(P) = \mu a P^2 \qquad (8)$$

과 같이 $\phi(P)$를 P의 증가함수로 두었다. 그 결과 다음 1계 선형 미분 방정식;

$$\frac{1}{P}\frac{dP}{dt} = \mu - \mu a P \qquad (9)$$

를 얻었다.

위의 미분방정식을 풀고 정돈하면 Logistic 곡선;

$$P_t = \frac{1}{a + b e^{-\mu t}} \qquad (10)$$

또는

$$P_t = \frac{1/a}{1 + b e^{-\mu t}/a} \qquad (11)$$

를 얻는다.

여기서 b는 적분상수이고, $1/a(=N)$은 소위 종국인구(Ultimate Population)으로서 시간 t가 무한히 흐를때 ($t \rightarrow \infty$) 인구크기 P의 극한치로 인구증가의 상한점근값(Upper Asymptote Value)를 얻게 되는 것이다. 그리고 인구 P의 하한점근치(Lower Asymptote value)는 영(零)이다. 또 Logistic곡선은 그 변곡점(變曲點)으로 $P_t = 1/2a$를 가지게 되며 이 점을 중심으로 S-자형의 점대칭곡선을 이룬다.

(b) Logistic곡선의 재발견

1920년 R. Pearl 과 Reed는 Verhulst와는 독립적으로 Logistic곡선식을 유도하였고, 그들은 1790년 부터 1910년 사이에 실시된 미국 Census 인구자료를 바탕으로하여 다음과 같이 미국인구의 Logistic곡선을 추정하였다.

$$P_t = \frac{197{,}273{,}000}{1 + 67.32\, e^{-0.0313t}} \tag{12}$$

(c) Logistic곡선식의 모수추정

이제부터 인구성장곡선을 위한 수리모형으로 단순로지스틱곡선(Simple Logistic Curve)를 선택하고 그것의 모수(Parameter)들을 추정하기위한 방법론을 소개하여 보기로 하겠다.

앞서의 (b)항 (11)식을 다음과 같이

$$P_t = \frac{k}{1 + e^{a+bt}} \tag{13}$$

로 다시쓰고 3-군법(3-Groups Method)에 의하여 (13)식의 모수 a, b 및 k 값들을 추정하기로 하겠다.

그러기 위하여 (13)식을 또 다시 다음과 같이 변형하기로 하자.

$$P_t = \frac{100{,}000}{a + uv^t} \tag{14}$$

여기서 $ak = 100{,}000$, $u = a\exp[a]$, $v = \exp[b]$, 이제

$$Q(t) = a + uv^t = \frac{100{,}000}{P_t} \tag{15}$$

로 두고, (15)식에 3-군법을 적용하여 u, v, a 등을 결정하는 방법을 소개하기로 한다.

<표-1> 3-군법의계산표

연도별 (t)	$Q(t)$ 값의 계산
$t=0$ $t=1$ $\vdots$ $t=n-1$	$a+u=100{,}000/P_0$ $a+uv=100{,}000/P_1$ $\vdots$ $a+uv^{n-1}=100{,}000/P_1$
$\sum(1)$: 제 1군합	$na+u(v^n-1)/(v-1)=100{,}000\sum_{t=0}^{n-1}P_{n-1}$
$t=n$ $t=n+1$ $\vdots$ $t=2n-1$	$a+uv^n=100{,}000/P_n$ $a+uv^{n+1}=100{,}000/P_{n+1}$ $\vdots$ $a+uv^{2n-1}=100{,}000/P_{2n-1}$
$\sum(2)$: 제 2군합	$na+uv^n(v^{n-1})/(v-1)=100{,}000\sum_{t=n}^{2n-1}P_t^{-1}$
$t=2n$ $t=2n+1$ $\vdots$ $t=3n-1$	$a+uv^{2n}=100{,}000/P_{2n}$ $a+uv^{2n+1}=100{,}000/P_{2n+1}$ $\vdots$ $a+uv^{3n-1}=100{,}000/P_{3n-1}$
$\sum(3)$: 제 3군합	$na+uv^{2n}(v^{n-1})/(v-1)=100{,}000\sum_{t=2n}^{3n-1}P_t^{-1}$

이제 <표-1>의 3-군법의 계산에서 제 2군합과 제 1군합의 차(差)와 제 3군합과 제 2군합의 차로 정의되는 정차(定差)들을 각각 $\Delta\sum(1)$ 과 $\Delta\sum(2)$로 쓰기로 하면

$$\Delta\sum(1)=\sum(2)-\sum(1)=u\frac{(v^n-1)^2}{v-1} \tag{16}$$

$$\Delta\sum(2)=\sum(3)-\sum(2)=uv^n\frac{(v^n-1)^2}{v-1} \tag{17}$$

과 같다. 따라서 (16)식과 (17)식에서

$$\frac{\Delta\sum(2)}{\Delta\sum(1)}=v^n \tag{18}$$

를 얻는다.

또 (16)식에서 $v=c$로 두고 u에 관해 풀면

$$u = \frac{\Delta\sum(1)(c-1)}{(c^n-1)} \tag{19}$$

와 같다. 그러므로 <표-1> 의 제 1군합을 a에 관해 풀고 (19)식의 u의 값을 대입하면

$$a = \frac{1}{n}\left\{\sum_{t=0}^{n-1}\frac{100{,}000}{P_t} - \frac{\Delta\sum(1)}{c^n-1}\right\} \tag{20}$$

와 같다.

고로 (18)식에서 얻어지는 $v=c$ 와 (19)식의 u의 값 (20)식의 a의 값들을 추정치로 하여 (14)식에 대입하면 구하려고 하는 Logistic곡선이 결정된다.

즉, 구하려고 하는 Logistic곡선 식:

$$\begin{aligned} P_t &= 100{,}000\{a + uc^t\}^{-1} \\ &= \frac{K}{\left[1 + \left(\frac{u \times c^t}{a}\right)\right]}, \end{aligned} \tag{21}$$

여기서 $K = 100{,}000 \times a^{-1}$를 얻는다.

(d) Logistic 곡선의 변곡점 계산:

앞서의 (a)항 (11)식에서

$$\mu = \lambda \quad , \quad \frac{1}{a} = N \quad , \quad \frac{b}{a} = k$$

로 두면 로지스틱 곡선의 표준형:

$$P_t = \frac{N}{1 + e^{\ln k - \lambda t}} \tag{22}$$

을 얻는다. 그리고, 이 곡선식에서 그 변곡점(Inflexion Point)를 구하여 보면 다음과 같다.

$$P_t'' = \frac{\lambda^2 N e^{\ln k - \lambda t}(e^{\ln k - \lambda t} - 1)}{(1 + e^{\ln k - \lambda t})^3} = 0$$

$$\Leftrightarrow \quad t = \frac{\ln k}{\lambda} \tag{23}$$

$$\Leftrightarrow \quad t = \frac{1}{2a} \quad . \tag{24}$$

3. 한국 인구성장곡선의 추정

이제부터 우리나라 인구성장 곡선을 추정하기위하여 그 수리모형 곡선으로서 2절의 (21)식의 단순 Logistic 곡선 (Simple Logistic Curve)을 선정하고 그 곡선의 모수들인 a, μ, c등의 추정치(Estimate)들을 구하여 보기로 하겠다. 이때 3-군법을 적용할 기본자료군으로서는 1949년부터 1990년까지 9회에 걸쳐 실시된 인구census 자료들을 이용하였다.

즉, 앞서의 서론에서도 언급하였듯이 1950년의 총인구에 관한 Census자료가 없으므로 1949년 Census 자료와 1955년 자료를 이용하여 1950년의 총인구를 선형보간(Linear Interpolation)하였으며, 1965년의 앙년추계인구(央年推計人口)를 1966년 Census인구를 대신하여 1950년 부터 1990년 인구Census 자료까지를 이용하여 5년 간격의 Census 자료로 수정 보완하였다.

<표-2> 우리나라 인구 Census 자료(단위:천명)

구 분	연도(Yrs)	변수 (t)	총 인 구	$Q(t)$	$\sum(i)$
제 1군 자료	1950 (5.1)	0	20,412	4.899	
	1955 (9.1)	1	21,526	4.646	$\sum(1)$ = 13.547
	1960(12.1)	2	24,989	4.002	
제 2군 자료	1965 (7.1)	3	28,705	3.484	
	1970(10.1)	4	31,466	3.178	$\sum(2)$ = 9.543
	1975(10.1)	5	34,707	2.881	
제 3군 자료	1980(11.1)	6	37,436	2.671	
	1985(11.1)	7	40,448	2.472	$\sum(3)$ = 7.441
	1990(11.1)	8	43,520	2.298	

① 모수 c의 추정치의 계산:

$$\Delta\sum(1) = \sum(2) - \sum(1) = -4.004$$

$$\Delta\sum(2) = \sum(3) - \sum(2) = -2.102$$

$$\therefore\ c^3 = \frac{\Delta\sum(2)}{\Delta\sum(1)} = 0.525$$

$$\therefore\ \hat{c} \approx 0.79144$$

② 모수 a의 추정치의 계산:

$$\hat{a} = \frac{1}{n}\left[Q(1) - \frac{\Delta\sum(1)}{c^n - 1}\right]$$
$$= \frac{1}{3}\left[13.457 - \frac{(-4.064)}{0.525 - 1}\right]$$
$$\approx 1.706$$

③ 모수 u의 추정치의 계산:

$$\hat{u} = \frac{\Delta\sum(1)(c - 1)}{(c^n - 1)^2}$$
$$= \frac{(-4.004)(0.79144 - 1)}{(0.475)^2}$$
$$\approx 3.701$$

그러므로 위 모수들의 추정치들을 (21)식에 대입하여 다음 인구성장곡선을 얻게 된다.

$$P_t = \frac{58.616}{1 + \left[\frac{3.701 \times (0.79144)^t}{1.706}\right]} = \frac{58.616}{1 + 2.169 \times (0.79144)^t} \qquad (22)$$

그리고 다시 (22)식을 단순 Logistic곡선의 표준형으로 바꿔서 주면 다음(23)식을 얻는다.

$$P_t = \frac{58.616}{1 + \exp\{\ln 2.169 - 0.2339\,t\}} \qquad (23)$$

4. 결과 및 결론

위 인구성장에 관한 (22)식 또는 (23)식에 의하여 각 인구Census구간 사이의 연도별 총인구를 계산하여 보면 <표-3>과 같다.

<표-3> 한국인구의 연도별 Census구간 인구 추계치 (단위:1,000명)

연 도	변 수(t)	Census인구	곡선추정인구	선형보간인구
1950	0.0	20,412	18,496.7	
1951	0.2	.	19,094.0	20,634.5
1952	0.4	.	19,701.1	20,857.6
1953	0.6	.	20,317.5	21,080.4
1954	0.8	.	20,942.9	21,303.2
1955	1.0	21,526	21,576.7	21,526.0
1956	1.2	.	22,218.4	22,218.6
1957	1.4	.	22,867.3	22,911.2
1958	1.6	.	23,523.0	23,603.8
1959	1.8	.	24,184.7	24,296.4
1960	2.0	24,989	24,851.9	24,989.0
1961	2.2	.	25,523.8	25,731.3
1962	2.4	.	26,199.8	26,475.4
1963	2.6	.	26,879.2	27,218.6
1964	2.8	.	27,561.2	27,916.8
1965	3.0	28,705	28,245.1	28,705.0
1966	3.2	.	28,930.2	29,257.2
1967	3.4	.	29,615.7	29,809.4
1968	3.6	.	30,300.8	30,316.6
1969	3.8	.	30,984.9	30,913.8
1970	4.0	31,466	31,667.1	31,466.0
1971	4.2	.	32,346.8	32,114.2
1972	4.4	.	33,023.2	32,762.4
1973	4.6	.	33,698.6	33,410.6
1974	4.8	.	34,363.3	34,058.8
1975	5.0	34,707	35,025.6	34,707.0
1976	5.2	.	35,681.9	35,252.8
1977	5.4	.	36,331.6	35,798.6
1978	5.6	.	36,974.6	36,344.4
1979	5.8	.	37,608.6	36,890.2
1980	6.0	37,436	38,234.9	37,436.0
1981	6.2	.	38,852.3	38,038.4
1982	6.4	.	39,460.4	38,640.8
1983	6.6	.	40,058.7	39,243.2
1984	6.8	.	40,646.8	39,845.6
1985	7.0	40,448	41,224.4	40,448.0
1986	7.2	.	41,791.1	41,062.4
1987	7.4	.	42,346.6	41,676.8
1988	7.6	.	42,890.7	42.291.2
1989	7.8	.	43,423.1	42,905.6
1990	8.0	43,520	43,943.6	43,520.0

다른한편, Logistic곡선 (22)식 또는 (23)식으로 1990년도 이후 2010년까지의 연도별 총인구의 장래추계치는 <표-4>와 같다.

<표-4> 한국인구의 연도별 장래 예측치(단위:1,000명)

연 도 별	변 수	장래예측인구
1991	8.2	44,452.2
1992	8.4	44,948.6
1993	8.6	45,432.7
1994	8.8	45,904.6
1995	9.0	46,364.1
1996	9.2	46,811.3
1997	9.4	47,246.1
1998	9.6	47,668.7
1999	9.8	48,079.1
2000	10.0	48,477.4
2001	10.2	48,863.7
2002	10.4	49,238.1
2003	10.6	49,600.7
2004	10.8	49,951.9
2005	11.0	50,291.6
2006	11.2	50,620.3
2007	11.4	50,937.7
2008	11.6	51.244.5
2009	11.8	51,540.7
2010	12.0	51,826.6

그리고 또 인구성장곡선 (23)식을 그래프로 나타내주고, 그 위에 Census인구를 점선으로 그려보면 <그림-1>과 같다.

한편, 그래프상에서 이론곡선과 실제 Census자료의 점들의 접근성을 볼때 Logistic곡선의 적합성(Fitness)을 인정할수있고, 특히 위 Logistic곡선의 결정계수(Coefficient of Determination)로서 실측치(Census人口)와 그에 대응하는 이론치(곡선추정치) 사이의 상관계수를 계산하였더니 그값이 $r = 0.998$과 같이 높은 상관관계를 보였다.

그러므로 2000년대의 총인구의 추정(예측)도 가능하겠으며, 인구증가 추세는 그 종국 총인구(Ultimate Total Population) 58,616천명을 초과 할 수 없겠음을 예견 할 수 있다.

끝으로 Logistic곡선 (22)식에 관하여 방정식:

$$\ln 2.169 - 0.2339\, t = 0$$

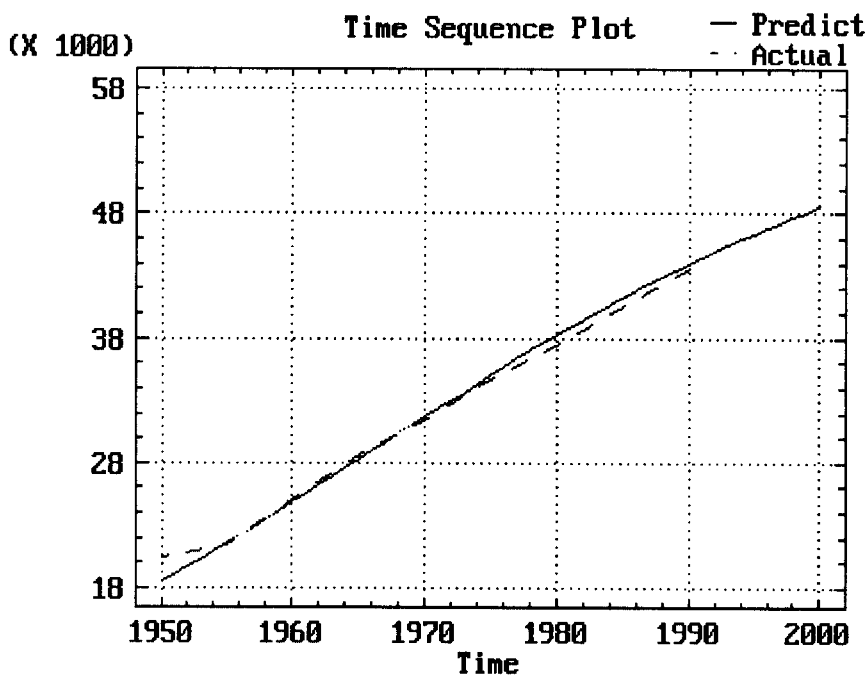
의 해를 구해 보면

$$t \approx 3.310$$

을 얻는다.

이때 t값에 해당하는 연도(年度)를 찾아보면 대략1967년도로서 그해의 총인구는 29,615천명이어서 총국인구의 반분점(半分點)이 된다. 특히 <표-3>의 제 5열의 선형보간법에 의한 연도별 총인구 추정치들을 제 4열의 곡선추정치들과 비교하여 보면 일반적으로 대략 1967년이전에는 보간인구가 곡선추계인구보다 높고, 그 이후는 곡선추정인구가 보간인구보다 높이 추정된다.

이와같은 반분점은 우리나라 인구가 1967년까지는 급성장(急成長)을 하다가 그 이후 점차로 완만한 성장을 하게되는 인구성장의 전환점(轉換點)을 의미한다고 볼 수 있다.



<그림-1> 한국인구 곡선과 실측치의 비교

참고문헌

[1] 崔鍾碩(1971), 우리나라 인구증가 추세에 따른 이론적경향선의 추정,『충남대 논문집』, 8-9.
[2] 山口喜一外(1989).『人口分析學』, 古今書院, 東京, 日本, 61-62.
[3] 岡崎陽一(1990).『人口統計學』, 古今書院, 東京, 日本, 141-179.
[4] Jaffe, A. J. (1951). *Handbook of Statistical methods for Demographers,* United States Government Printing Office, Washington D. C. 211 - 218.
[5] Min, T. (1991). *Korea Statistical Year Book,* Vol. 38, National Statistical Office, Republic of Korea, 35.
[6] Minoru Tachi (1963).『人口分析の方法』, 古今書院, 東京, 日本, 54, 87-90
[7] Spieglman, M. (1955). *Introduction to Demography,* Havard, 407 - 408.

On the Estimation of Parameters for the Growth Curve of the Korean Population[3)]

Ja-Heung, Koo[4)]

Abstract

The purpose of this research is to obtain a Simple Logistic Curve for the curve fitting of Korean total Population. Based on the population census data from 1949 to 1990, the parameters are estimated by 3-group method. As the results, intercensal populations of Korea from 1950 to 1990 are estimated, and Korean total populations from 1991 to 2010 A.D. are projected. And we also can suggest the upper asymptote 58,616 thousands of Korean total population.

3) This research was supported by Inha University Research Grant, 1992.
4) Department of Statistics, University of Inha, 253 Yunghyun-Dong, Namku, Inchon, 402-751, KOREA.