

9月27日は十五夜。ススキや月見だんごを飾るという
習わしもうすれているようだ。

団地のベランダから見る月は、いつも黄色くよごれて
いる。あれは、宇宙飛行士が月の砂をかきまわしてしま
ったせいだろう。

今月のおもな行事

- 2～3日 常陸太田市統計大会（栃木県）
- 8日 茨城県統計グラフコンクール締切
- 8～9日 工業統計調査主管課長会議（長野県）
関東ブロック県単統計研究会（静岡県）
- 12～13日 農業基本調査協議会（千葉県）
- 16日 消費者動向調査員会議（自治会館）
- 27～28日 広報連絡研修会議（旭村・いこいの村瀬沼）
- 28～30日 図書館大会（大阪府）

「回帰」と「相関」 —関係分析の方法—

「カイキ」と「ソーカン」と書きますと、何やら誤解をまねきそうですが、統計の世界では漢字で「回帰」と「相関」と書きます。これは回帰分析と相関分析のことです。両方とも「2変量間の関係を統計的に分析する」ための方法です。

回帰分析

回帰分析は、「2変量のうち1つを独立変数あるいは原因変数と考え、この独立変数を変化させるときに生ずる効果を他の変量——従属変数あるいは結果変数の変化で観察しようとする分析」をいいます。さて、これだけでは良くわかりません。そこで統計いばらき6月号の3ページを開いて下さい。「最小2乗法」について掲載されているからです。良くみると、これは年次と観察値という2つの変量の関係が題材になっているようです。

実は、これは時系列データの分析で、年次ごとの傾向を調べたものです。これを傾向直線といって、回帰分析によって表わされる回帰直線（変量間の関係を表現する数式で表わされた回帰線の中で、特に $y = a + bx$ で現わされた直線をいう）の一種なのです。傾向直線は年次と観察値との関係で、回帰直線（回帰線）が観察値と観察値との関係であることと比べれば、年次も観察値の一種と見ているという訳になります。

今、回帰直線を

$$y = a + bx \quad (a, b \text{ は定数})$$

の数式で現わしてみましょう。この数式の係数 a, b の計算に最小2乗法を利用するという事は学習済みですね。 y は従属変数、 x は独立変数で、 a, b の値がわかっているならば、 x の任意の変化に対応して y の値が定まります。その値を図示して直線で結べば良いのです。このように直線で表現されている関係を直線回帰といいます。

ところで、一般に回帰線は直線に限定されている訳ではありません。2次式、3次式などの多項式や指数曲線などが最小2乗法で当てはめられる場合もあります。その際の

計算はかなりやっかいで、コンピューターを使わなければならない場合もあります。

回帰分析では、独立変数を自由かつ正確にコントロールできるということが必要な前提であって、実験による観察結果の分析に特に適しており、生物学や自然科学ではきわめて重要な分析方法として利用されています。社会現象や経済現象では条件を自由にコントロールできる実験が困難ですから、厳密な意味での回帰分析の応用は困難です。しかし、実際にはしばしば応用されています。これは、正確に言えば近似的あるいは代用的な応用とみるべき場合が多いのです。社会現象、特に経済現象の分析で回帰分析の方法が好んで用いられているのは、経済変数間の関係を数学の方程式、たとえば直線やその他の多項式などの形で表わすことができれば、それを推定あるいは予測のための経験式として利用することが可能だからなのです。非実験的なデータに対する回帰分析の応用の制限をどのようにして解決するか、ということが、回帰分析を経済学に応用する際の課題となっているのです。

参考までに、2次、3次などの式によって現わされる回帰線（このような数式は曲線を表わすので、このような関係を曲線回帰といいます）として把握した方が良いと考えられている例をあげましょう。

(例)

①年齢と賃金 ②年齢と食物摂取量 ③窒素投下量と作物収量 etc

なん次の多項式をとったら良いかは、実際にいろいろ当てはめて試算してみた上で判断する以外に手はなさそうです。

相関分析

つぎに相関分析をみてみましょう。これは、「2変量が、無差別に選ばれた観察単位の相対応する変量である場合に、その関係の分析をいう。1つの変量を他の説明要因として規定しない」ということです。これは要するに、回帰分析

伊藤 宰

の場合は、2つの変量を独立変数 x と従属変数 y とに定義したのに対し、相関分析の場合にはこれと同じでも良いし、 x が従属変数、 y が独立変数となっても良いということを意味します。つまり回帰分析の時ほどには、 x と y との関連がわかっていない場合のことなのです。このように、回帰分析と相関分析は表裏一体となっています。ですから、回帰線は相関分析の場合にも定義することができます。この場合には、2つの変数 x 、 y のいずれかを独立変数と定めてよいのですから、直線回帰の場合には、

$$y = a + bx \quad \text{または} \quad x = a' + b'y$$

(a, b, a', b' は定数)

と、2つの回帰直線が定義できます。もちろん1次式だけでなく、2次式、3次式などの回帰曲線もそれぞれを独立変数、従属変数として定義できます。

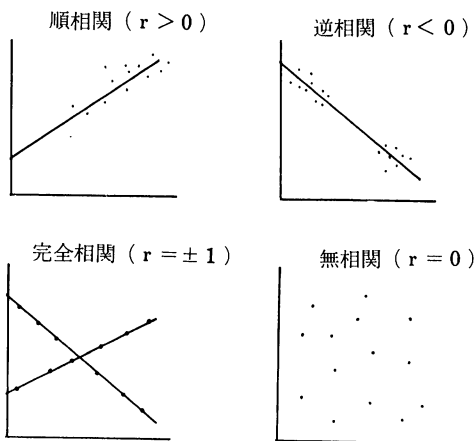
同一のデータに対して、2つの異なった回帰線が存在することは意外に感じるでしょうが、変量の扱い方のちがいにより別のものが現われるのはむしろ当然でしょう。時系列データ（6月号3ページ参照）の場合には、独立変数として扱われるのが常に時間であるため、 y の x に対する回帰線のみが傾向線として考えられている訳で、これは特殊な例なのです。

さて、直線回帰の場合、変数 x 、 y 間の相互関係の強さを測定する方法があります。

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

$$= \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}}$$

というむずかしそうな式で現わされるのですが、この r を相関係数といって、 $-1 \leq r \leq 1$ 、つまり、 -1 から 1 までの間をウロウロします。この値によって、



という関係を示しています。これは、回帰線が直線であることを前提としていますので、曲線回帰の場合には相関比という考えを使います。

このほか、多変数回帰分析（多元回帰分析）、重相関、偏相関などさまざまな概念があります。

では、今まで出てきた考え方を応用してみましょう。王選手のホームラン本数と巨人軍の勝率との関係についてです。

(例) 王選手のホームラン本数と巨人軍の勝率(過去10年間)

年次	本数 x	勝率 y	xy	x^2	y^2
42	47	0.646	30.362	2,209	0.417316
43	49	0.592	29.008	2,401	0.350464
44	44	0.589	25.916	1,936	0.346921
45	47	0.627	29.469	2,209	0.393129
46	39	0.574	22.386	1,521	0.329476
47	48	0.587	28.176	2,304	0.344569
48	51	0.524	26.724	2,601	0.274576
49	49	0.587	28.763	2,401	0.344569
50	33	0.382	12.606	1,089	0.145924
51	49	0.628	30.772	2,401	0.394384
Σ	456	5.736	264.182	21,072	3.341328

$\bar{x} = 45.6 \quad \bar{y} = 0.5736 \quad (\bar{x}, \bar{y} \text{ は } x, y \text{ の平均値})$

●シリーズ「統計」

1) y を従属変数, x を独立変数とする。($y = a + bx$ の型)

$$\text{正規方程式} \begin{cases} \sum y = na + b\sum x \\ \sum xy = a\sum x + b\sum x^2 \end{cases} \text{に代入して}$$

$$5.736 = 10a + 456b \quad \text{---①}$$

$$264.182 = 456a + 21,072b \quad \text{---②}$$

$$\text{①} \times 45.6$$

$$261.5616 = 456a + 20,793.6b \quad \text{---①'}$$

$$\text{②} - \text{①'}$$

$$2.6204 = 278.4b$$

$$\therefore b = 0.00941235$$

①に代入して

$$a = 0.14439684$$

したがって, $y = 0.14439684 + 0.00941235x$

$$\approx 0.144 + 0.009x$$

2) y を独立変数, x を従属変数とする。($x = a' + b'y$ の型)

$$\text{正規方程式} \begin{cases} \sum x = na' + b'\sum y \\ \sum xy = a'\sum y + b'\sum y^2 \end{cases} \text{に代入して}$$

$$456 = 10a' + 5.736b' \quad \text{---①}$$

$$264.182 = 5.736a' + 3.341328b' \quad \text{---②}$$

$$\text{②} \times 10 - \text{①} \times 5.736$$

$$26.204 = 0.511584b'$$

$$\therefore b' = 51.221304$$

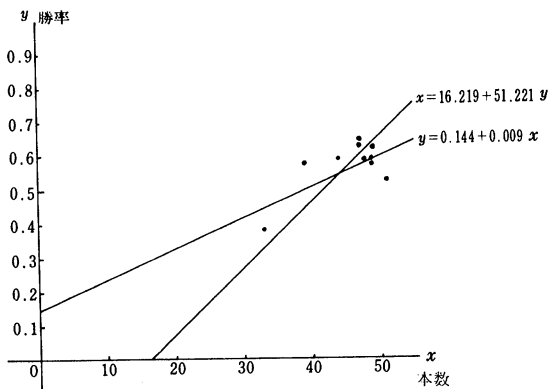
①に代入

$$a' = 16.219461$$

したがって, $x = 16.219461 + 51.221304y$

$$\approx 16.219 + 51.221y$$

以上の回帰直線を図示してみましょう。



ホームラン本数によって, 勝率がどう変化するかをみるのが, ホームラン本数 x を独立変数, 勝率 y を従属変数とする $y = a + bx$ の型です。逆に勝率によってホームラン本数の変化をみるのが $x = a' + b'y$ の型です。

次に相関係数を計算してみましょう。

$$r = \frac{264.182 - 10 \times 45.6 \times 0.5736}{\sqrt{(21,072 - 10 \times 2,079.36)(3.341328 - 10 \times 0.3290169)}}$$

$$= \frac{2.6204}{\sqrt{14.242665}}$$

$$= \frac{2.6204}{3.7739455}$$

$$= 0.6943396$$

$$\approx 0.694$$

r は 0.694 となりました。 $r > 0$ ですから順相関, つまり王選手のホームラン本数と勝率とは, お互いに増加するときは増加, 減少するときは減少するという関係にあることがわかりました。やはり王選手はすばらしいバッターといえましょう。

(県統計課 企画調整係)

茨城県の人口に関する考察

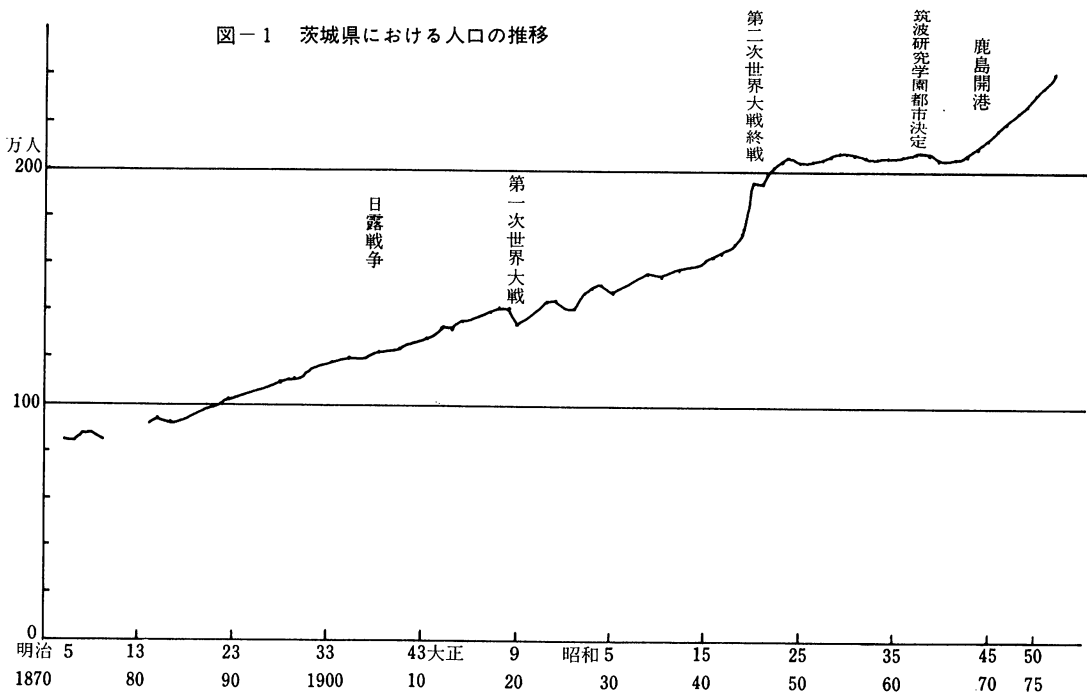
堀 口 友 一

茨城県の地域性は歴史的に大きく変わっている。江戸時代 260年間の水戸藩は、徳川ご三家の一として天下に名をなした雄藩であった。時代が移り、明治になってから昭和前期（終戦まで）に至るまでの茨城県は、農業県という性格に終始した。明治中期におけるわが国の産業革命にも、近代工業の立地条件に恵まれず、食糧生産には貢献しながらも、京浜から北九州に至る諸都府県に比らべると、全般的に発展への立ち遅れを示していた。昭和30年以後、わが国の経済高度成長と政府および県当局の開発施策は、本県発展への性格を大きく変えようとしている。東海村の日本原子力研究所、鹿島開発、筑波研究学園都市の建設などは首都圏整備とともに、農業県という基盤の上に新たな性格を指向する。このような本県の地域性的変化が、本県の人口とどのような関係にあるかを考察するのが、本稿の意図である。

人口の推移 水戸藩では寛永15～16年（1638～39）初代藩主頼房のとき人馬改を行い、年貢、夫役を確保するため

に領内の戸口、牛馬を調査したのが初めて、その後たびたび人別改が行われている。これらの人別改によると、元禄10年（1697）に 276,926人で、女子 100人に対して男子120.4人の割合を示す。延享4年（1744）には275,820人、文政11年（1828）には231,766人に減少し、元治元年（1864）に120.4人の割合を示す。延享4年（1744）には275,820人、文政11年（1828）には231,766人に減少し、元治元年（1864）には274,908人で、167年前の元禄10年の人口に比らべて約2,000人減少している。

明治4年（1871）7月廃藩置県が行われ、明治8年（1875）5月茨城県は新治県その他を合わせて、今日の茨城県が誕生した。明治9年（1876）1月1日の人口は 857,882人で、明治22年（1889）に 100万を越えた。明治・大正年間における人口増加は比較的緩慢で、昭和22年（1947）に200万を越え、昭和50年の国勢調査で2,342,173人（男1,159,772、女1,182,401）、昭和52年6月1日現在（茨城県統計課）で2,402,322（男1,192,166、女1,210,156）を示している。第

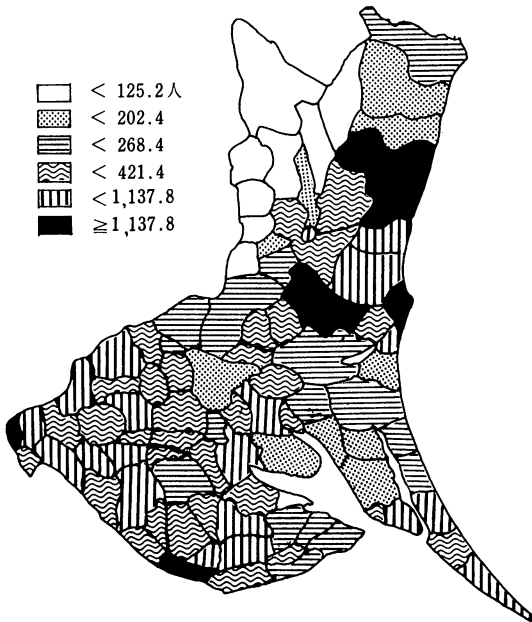


統計の窓「論壇」

1図によれば、明治のころは緩やかな増加を示し、大正9年（1920）の急減後約10年間増減を繰り返し、第2次大戦終結の昭和20年（1945）に驚異的な増加がみられる。終戦後の昭和22年より同43年までは著しい増減はみられなかったが、昭和44年以後急激な増加を示している。

人口分布 人口分布の状態は人間活動の総和として把握される。しかし人口分布の地域把握は人口の総数よりも、密度によることがより合理的である。昭和50年国勢調査における本県の人口密度は1km²につき385人で、全国の都道府県に比較すると13位に当る。関東地方では東京5,443人、神奈川2,383人、埼玉1,269人、千葉813人で、いずれもわが国における稠密地域を形成し、茨城県はこれらに比し著しく低下するが、群馬の276人、栃木の265人より高い。

図-2 茨城県における市町村別人口密度（階級区分は棄却限界公式による）



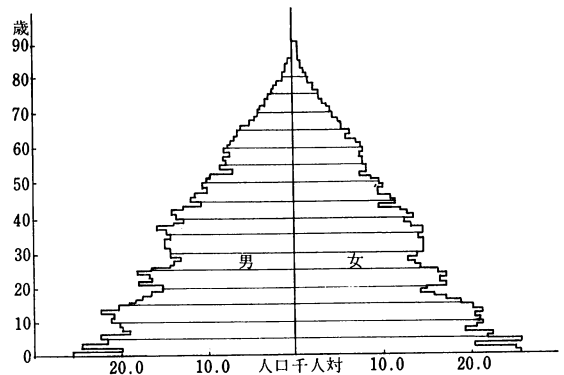
昭和51年茨城県統計年鑑により、県内92市町村の人口密度（昭、49）を棄却限界の公式から算出した階級区分の分布図が第2図である。1方キロ当たり1,137.8人以上の地域は日立、那珂湊、水戸、古河、取手の5市である。421.4～1,137.8人の地域は勝田、石岡、土浦、竜ヶ崎、下館、結城、下妻、水海道の8市と、東海、大洗、鹿島、波崎、潮来、

藤代、牛久、石下、猿島、境の10ヶ町村である。これらの稠密地域に対し、125.2～202.4人の高萩、桂、十王、金砂郷、旭、北浦、麻生、玉造、出島、八郷の市町村と、125.2人以下を示す御前山、七会、緒川、美和、山方、大子、水府、里美等の北部山地の町村は低密地域を構成している。

ちなみに、本県の人口を全国と比較した場合、昭和50年10月1日の国勢調査結果によると、人口総数では都道府県別で14位にあたる。

人口構造 年令別・性別人口構造図によると、明治・大正時代の本県の形は先鋭なピラミッド型を示していた（第3図）。この型は一般に発展途上国にみられるタイプで、当

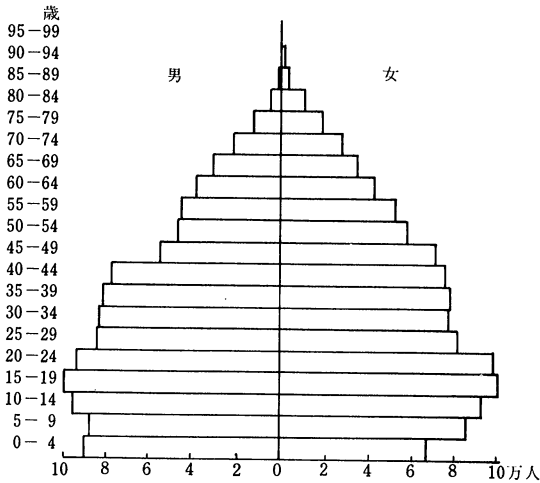
図-3 性別・年令別人口構成（明治22年）



時わが国の人口構造と同様で、乳幼児あるいは若年令層に死亡者が多いために表われる構造で、文化水準の低い結果による。昭和45年（1970）国勢調査結果による5歳別年令別・性別人口構造図（第4図）は壺型を示し、先進国型に移行している。この人口構成をさらに検討してみると、年令階層別では15～19歳が最も多く男女とも4.6%、次いで20～24歳が男4.4、女4.5%を示し、10～14歳、0～4歳の順となっている。45歳を越えると3%以下となり、50歳から高年令に移るにしたがって男女とも激減する。

人口構成の性別について比較すると、0歳人口は男子18,900に対し、女子18,233で男子が667人多い。男子人口が女子人口より多いのは20～24歳を除いて、40～44歳までで、45歳を越えると女子人口が多くなる。しかも女子対男子人口の比は、高令化するにしたがって、男子人口が激減

図-4 性別・年齢別人口構成（昭和45年）

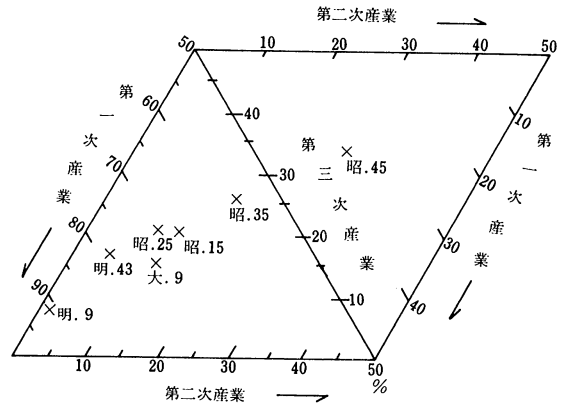


する。たとえば、50～54歳で女子 100に対して男子80.7であるが、70～74歳では男子は79.2となり、80～84歳は54.3、90～94歳は29.1となる。60歳以上の女子人口は138,991で、男子人口は 112,630、女子が26,361人多く、その比は女子人口 100に対して、男子81.0となる。

江戸時代における水戸藩の性別人口構成は、元禄10年(1697)女子 100人に対し、男子120.4、延享4年(1747)男子124.3、文政11年(1828)108.8で、男子人口の多いことを示している。この傾向は明治35年(1902)までつづき、明治9年(1876)には女子100に対して男子102.3、明治35年には101.1となっている。しかし、明治36年より男女人口の比は逆転している。同年に女子 100に対し、男99.7、明治43年(1910)に99.5、大正9年(1920)に88.9、昭和10年(1935)97.9となり、昭和16年より男子人口が急激に低下し、昭和19年(1944)には87.0の最低を示している。

本県における産業別人口構成をみると(第5図)、明治9年には第1次産業人口(農林水産業)は90.7%であるのに対し、第2次産業人口(鉱工業、建設業)はわずかに1.4%、第3次産業(サービス業)は7.9%である。明治43年(1910)においても第1次産業は79.6%を示し、大正9年に第2次産業が10%台に達している。昭和25年(1950)の構成をみても、第2次産業は大正9年に比してほとんど変化はなく、第1次産業の減少した3.7%分が、第3次産業に

図-5 茨城県における産業人口比率



プラスしているに過ぎない。本県の産業別人口構成に著しい変化がみられるようになったのは昭和30年(1955)以後で、昭和35年には第1次産業人口56.4%、第2次産業人口17.3%、第3次産業人口26.3%となり、昭和45年には第1次産業人口37.0%、第2次産業人口28.4%、第3次産業人口34.6%を示している。昭和45年のわが国における産業別人口比率と比較した場合、第1次産業においては17.6%高く、第2次産業、第3次産業においてそれぞれ、5.6%、12.0%低くなっている。

人口動態 昭和50年国勢調査による本県の人口は2,378,220人で、うち男1,178,688、女 1,199,532人である。昭和51年茨城県統計年鑑によりその増減状況を見ると、昭和50年の出生は42,474で、全人口 2,290,214に対し、1.85%、死亡は17,090で0.74%、差引自然増加数は25,384で、その比率は1.10%である。さらに同年における人口の社会増加についてみると、県外からの転入は74,389人、県外への転出は59,048人で、本県から転出する者より、他県より転入する者の数が15,048人多くなっている。昭和50年における転入、転出者を都道府県別にみると、転入の最も多いのが東京都で23,891人、これに次いで千葉県10,983、神奈川県7,984、埼玉県6,953、栃木県3,903、福島県3,012、群馬県790等の順を示している。県外への転出の内訳は東京都が最も多く18,406、千葉県9,998、神奈川県6,033、埼玉県5,624、栃木県3,188、福島県2,653、群馬県770、その他の諸県12,669となっている。このようにこの年度における社会増は15,048人で、その比率は0.65%にあたる。

■統計の窓「論壇」

昭和50年における市町村別人口の自然増減をみると、自然増加率1.5%以上を示すのは鹿島2.24%、勝田2.0、取手1.89、千代田1.88、日立、東海1.76、総和1.68、神栖1.58水戸1.51の市町村である。増加率の0.3%以下の低率を示すのは、金砂郷、緒川、美和、山方、玉里、瓜連の諸町村で、減少を示しているのは、桂、御前山、水府、里美等北部山地諸村と、新利根、桜川、美浦の県南の3か村である。

昭和51年茨城県統計年鑑より算出される市町村別人口の社会増加についてみると、増加率の最も高いのは桜村の13.1%で、鹿島町の6.7%がこれに次ぎ、取手の6.2%、牛久、藤代の4.5%、千代田の4.0%、守谷の3.1%、伊奈の3.0%の順となっている。同年における社会増の実人数で最も多いのは、取手の3,275人で、鹿島の2,492、桜村の1,953、牛久の1,254の順を示している。転入人口より転出の多い市町村、すなわち人口の社会減を示す市町村は、日立の646人が最も多く、大子の291、勝田の239、北茨城の186、常陸太田の126の順となっている。

以上を総合して人口増加率の最も高いのは、桜村の13.7%で、これに次いで鹿島の8.7%、取手の7.9%、牛久の7.8%、藤代、千代田の5.8%、守谷の4.1%、神栖3.2%の順となっている。以前著しい人口の伸び率を示していた日立、勝田は社会減のため増加率は低い。北部山地の市町村はいずれも人口減あるいは停滞を示している。

人口の県外流動についてみると、昭和45年における通勤通学者数は60,435人で、通勤者49,246、通学者11,189人である。通勤は東京都が28,734(58.3%)で、千葉県9,382(19.1%)、栃木県5,165(10.5%)、埼玉県3,657(7.4%)、神奈川県1,112(2.3%)、福島県783(1.6%)の順となっている。通学者も東京都が最も多く6,079人で、全通学者数の54.3%を占めている。次いで千葉県が2,211(19.8%)、栃木県1,245(11.1%)、埼玉県1,067(9.5%)の順を示す。流出人口の最も多いのは県南地域で、本県流出人口全体の52.7%を占め、県西地域が33.6%を示す。

人口構成に影響する諸要因、自然的要因として性別人口構成の問題がある。男女別人口を比較すると、出生数すなわち0歳人口は男子が多い。このことは本県だけに限らず人口に関する一般的事象で、性遺伝に関する問題である。

性遺伝については、1923年ベルギーのペインターがXY説を提唱している。各国の人口統計において、女性より男性が多く生まれるのは、授精競争において男子の精子がまさっているか、あるいは移動速度が早いためであると指摘している。しかし、人口全体として女子人口の多いことは、上述のように30~40歳以上の女子人口が多いことによるもので、女子の平均寿命が男子のそれより長い結果による。

本県における人口密度の地域構成が県南および県西に高く、県北の特に北西部に低いことは、平野に居住が密で、山地に疎である地形的要因である。

本県の人口構成に及ぼす社会的要因についてみると、明治以後の本県の人口推移において、その増加は最近まで比較的低迷していた。大正9~14年の増加率は4.3%(全国平均6.8%)、昭和5~10年、4.2%(全国平均7.5%)、昭和10~15年、4.6%(全国平均5.6%)で、全国平均より低率である。終戦後の増加率もきわめて低く、昭和25~30年1.2%(全国平均7.3%)、30~35年は0.8%減(全国平均4.6%増)、35~40年は0.4%(全国平均5.2%)を示している。このような増加の低迷状態は、自然増は順調を示すにもかかわらず、社会動態において県外転出が多いためである。これは本県の農業の性格が強く、京浜をはじめとする工業地帯に人口が吸収されたからである。

人口の社会的増加が著しくなったのは昭和44年以後で、昭和44年の転入73,215人、転出62,242人で10,973人の増加を示し、19,347の自然増と合せて30,320の増加、その率は1.4%となっている。この要因は鹿島開発をはじめとする県南の開発による人口増加の結果である。昭和50年に著しい人口増を示した鹿島、神栖両町と千代田村は工業開発に起因し、増加率13.7%の驚異的な数字を示した桜村の場合は研究学園都市の建設による。

人口密度が交通条件に緊密な関係を有することは今さら述べるまでもない。東京を中心とする首都圏の放射型交通幹線の一環をなす東北線、国道4号線に沿って高密度の古河市があり、常磐線、国道6号線に沿って取手、土浦、石岡、水戸、勝田、日立等の稠密地域が分布する。

メトロポールとしての東京の巨大都市化は、戦後の経済成長と交通機関の整備に伴う時間距離の短縮、首都圏整備

