

제2장

출생신고자료를 활용한 출생방정식 도출 연구

오진호

I. 서론

- 인구동태통계(Vital Statistics)는 어떤 정해진 기간에 관찰한 인구변동 요인의 변화를 의미하며, 혼인, 출생, 사망, 사산, 이동 등 인구의 구조나 성장에 의미있게 발생하는 사건을 정해진 기간에 걸쳐서 집계한 것을 말함
 - 이중 출생통계(Live births)는 28개 항목으로 구성된 출생신고서를 바탕으로 집계되며, 주요 항목으로는 출생자의 주소, 부·모연령, 부·모직업, 부·모교육정도, 동거기간, 임신주수, 다태아여부(단태아, 쌍태아, 삼태아 이상으로 구분), 출생아 체중, 모총출생아수(birth order)임
- 출생에 관련된 기존 연구는 여러 영향변수들에 대해 차별출산력을 규명한 것이 대부분임
 - 차별출산력은 출산력 감소의 현황과 인과구조에 관한 분석에 기반을 둔 것으로, 정책 수립을 위해 사회적 특성을 가진 집단에서 어떠한 속도로 출산력이 감소하고 있는지를 규명하는 것임
 - 이에 대한 연구는 인구학, 사회학, 경제학 등 여러 분야에서 접근할 수 있는데, 소득을 중심으로 하는 경제학적 접근이 주요한 패러다임으로 자리잡고 있음(Robinson, 1997)



- 하지만 소득이 출산에 미치는 영향력은 고정적이지 않고(변동을 내포) 지역·시기별로 다르며, 남성·여성임금이 출산에 미치는 영향이 국가·시기·출생아 순위별로 다름(Sleebos, 2003)
- 본 연구는 기존 연구들의 차별출산력의 영향요인을 규명하는 것도 중요하지만, 출생신고자료의 변수별 추세와 특징을 규명하고자 함
 - 출생신고자료를 종합하여 볼 수 있는 관계식을 도출한다면, 본 자료의 또 다른 이면을 볼 수 있다는 것에서 시도하였음
 - 이에 출생신고자료의 변수를 통합적으로 살펴 볼 수 있는 출생방정식 (birth equation)을 도출하고 의미를 찾고자 함
 - 기존연구는 출산력에 관련된 영향력 변수를 가족보건실태조사, 노동패널조사, 인구총조사 등의 자료에서 통계적인 분석으로 판별하는 것에 주안점을 둔 것임
 - 본 연구는 인구동태통계의 출생신고자료를 활용하여 모층출생아수와 다태아여부의 출생방정식의 연도별 패턴¹⁾과 이들의 관계규명을 도출하는 것이 주된 차이점임

II. 출생방정식 도출과 비교

1. 출생방정식 도출

- 출생방정식을 정의하기 이전에 인구방정식(vital equation)²⁾의 개념 정립이 필요
 - 인구방정식은 어느 지역의 인구성장을 결정하는 출생 및 전입과

1) 본 연구에서는 출생신고자료의 시계열 공통항목 통일성의 한계로 2007년~2012년 자료를 활용한다. 출생방정식의 도출에 앞서 출생신고자료의 기초통계량 분석결과는 [부록]을 참조

2) 자연증가와 사회증가의 합으로 표시되는 인구방정식은 다음과 같다.
 인구성장 = (출생-사망)+(전입-전출) = 자연증가 + 사회증가(또는 순이동)

인구성장을 방해하는 사망 및 전출의 4가지 요인으로 이루어지는데, 여기서 출생과 사망은 재생산 요인이라 부르며, 출생과 사망의 차이를 자연증가, 전입과 전출의 차이를 사회증가 또는 순이동이라고 함 (통계청, 2011)

- 따라서 인구방정식과 유사하게 출생방정식은 출생에 관련된 11개 세부 변수들의 연도별 관계 규명이라 할 수 있음(<표 2-1> 참조)

<표 2-1> 출생방정식 도출 활용변수

구분	사용변수
종속변수	다태아여부(Multiple births status), 모총출생아수(Birth order)
독립변수	부·모연령(Age of father and mother) 임신주수(Period of pregnancy) 다태아출산순위(Multiple births by birth order) 동거기간(Duration of cohabitation) 출생아체중(Birth weight) 출생장소(Live births by place of delivery) 부·모교육정도(Educational attainment of father and mother)

주: 소득대신 부·모연령과 부·모교육정도를 대리변수로 활용

- 인구방정식은 출생과 사망의 자연증가와 전입과 전출의 사회증가로 이루어진 인구동태방정식이라면 출생방정식(Birth Equation)은 출생항목들 간의 유기적인 관계식임
- 이런 함수식을 연도별로 도출해 본다면 어떠한 변수가 종속변수에 영향력이 높은지와 동향을 파악할 수 있음([그림 2-1], 식(2-1)참조)

[그림 2-1] 출생신고자료 주요항목 함수 관계식

부·모연령, 임신주수, 모총출생아수, 다태아출산순위, 동거기간, 출생장소, 다태아여부, 부·모 교육정도

종속변수

=

F(독립변수) + e

주 : 'F'는 선형함수를 의미하며, e는 오차를 의미한다.

- 본 연구에서는 아래 식처럼 모총출생아수 및 다태아여부를 종속변수로 간주하고 독립변수들의 연도별 변수관계를 도출함

$$\begin{aligned} \text{모총출생아수} &= F\left\{\begin{array}{l} \text{부(모)연령, 임신주수, 다태아출산순위, 동거기간} \\ \text{출생장소, 다태아여부, 부(모)교육정도} \end{array}\right\} + e \\ \text{다태아여부} &= F\left\{\begin{array}{l} \text{부(모)연령, 임신주수, 다태아출산순위, 동거기간} \\ \text{출생장소, 모총출생아수, 부(모)교육정도} \end{array}\right\} + e \end{aligned} \quad (2-1)$$

- 출생방정식의 최적모형을 도출하기 위해 아래 세 가지 과정과 다항로지스틱모형을 활용함
 - 과정1) 변수들의 단위차(세, 주, 년, kg 등)을 보정하기 위해 변수 표준화를 실시
 - 과정2) 범주형 변수의 영향을 알아보기 위해 더미변수(Dummy variable)를 고려
 - 출생장소는 자택, 병원, 기타이므로 2개의 더미변수, 다태아여부는 단태아, 쌍태아, 삼태아 이상이므로 2개 더미변수, 부·모교육정도는 무학, 초졸, 중졸, 고졸, 대졸, 대학원 이상으로 구분되므로 5개의 더미변수를 생성함



- 과정3) 다중회귀분석에서 가장 빈번히 발생하는 다중공선성³⁾을 제거하기 위해서 변수선택(Variable Selection)을 적용
- 다태아여부처럼 세 가지 범주 이상인 경우는 아래와 같이 다항로지스틱 회귀모형(Multinomial Logistic Regression)을 활용하여 분석함

$$\begin{aligned} \log\left[\frac{p(Y=\text{삼태아이상})}{P(Y=\text{단태아})}\right] &= \beta_{10} + \beta_{11}x_1 + \dots + \beta_{1p}x_p + e \\ \log\left[\frac{p(Y=\text{쌍태아})}{P(Y=\text{단태아})}\right] &= \beta_{20} + \beta_{21}x_1 + \dots + \beta_{2p}x_p + e \end{aligned} \quad (2-2)$$

$$\log\left[\frac{p(Y=\text{삼태아이상})}{P(Y=\text{쌍태아})}\right] = \log\left[\frac{p(Y=\text{삼태아이상})}{P(Y=\text{단태아})}\right] - \log\left[\frac{p(Y=\text{쌍태아})}{P(Y=\text{단태아})}\right]$$

여기서, 참조범주(reference category)는 단태아

- 일반적으로 다항로지스틱에서 상대적인 변수 영향력은 오즈비(odds ratio)로 파악하는데, 독립변수의 계수값이 0.5이면 오즈비는 1.648 [exp(0.05)]임
 - 1.648의 의미는 기준범주(단태아)가 다른 범주(쌍태아 이상)에 비해 해당변수(부연령, 다태아 출산순위 등)가 높은 값을 가질 확률이 1.6배 정도 높다는 것을 의미
- 모형적합을 위해 다항로지스틱의 전진선택법을 적용하여 최종모형을 우도비검정에 의해 선택함
- 이상 논의된 다중회귀분석과 다항로지스틱 모형은 식(2-3)~식(2-5)에 제시

3) 다중회귀분석에서 독립변수들 간의 선형관계가 발생하는 경우가 있는데 이를 다중공선성문제라 하고 추정된 회귀계수를 편이 추정되게 한다. 이를 진단하는 통계적인 방법으로 VIF지수가 있으며 이를 해결하는 방법으로 변수선택(variable selection), ridge regression, PCR 등이 있다. 또한 사회과학에서는 회귀분석의 오차항의 가정(등분산성, 독립성)을 따르지 않는 경우가 많으므로 일반화회귀분석(GLM)을 활용하여 해결한다. 본 연구에서는 변수선택과 GLM 분석을 모두 실시하여 결과를 비교 제시한다. 그 결과 유사한 회귀계수값과 표준오차가 도출되어, 본 연구에서는 회귀분석결과를 제시한다. 비교 결과는 [부록]의 <부표>를 참조



$$\begin{aligned}
 y_{\text{모충출생아수}_i} &= \beta_0 + \beta_1 \text{부연령}_i + \beta_2 \text{모연령}_i + \beta_3 \text{임신주수}_i + \beta_4 \text{동거기간}_i \\
 &+ \beta_5 \text{다태아출산순위}_i + \beta_6 I_{11}(\text{출생장소} = \text{주택}) + \beta_7 I_{12}(\text{출생장소} = \text{병원}) \\
 &+ \beta_7 I_{21}(\text{다태아여부} = \text{단태아}) + \beta_8 I_{22}(\text{다태아여부} = \text{쌍태아}) \\
 &+ \beta_9 I_{31}(\text{부교육} = \text{무학}) + \beta_{10} I_{32}(\text{부교육} = \text{초졸}) + \beta_{11} I_{33}(\text{부교육} = \text{중졸}) \\
 &+ \beta_{12} I_{34}(\text{부교육} = \text{고졸}) + \beta_{13} I_{35}(\text{부교육} = \text{대학}) \\
 &+ \beta_{14} I_{41}(\text{모교육} = \text{무학}) + \beta_{15} I_{42}(\text{모교육} = \text{초졸}) + \beta_{16} I_{43}(\text{모교육} = \text{중졸}) \\
 &+ \beta_{17} I_{44}(\text{모교육} = \text{고졸}) + \beta_{18} I_{45}(\text{모교육} = \text{대학}) + e_i
 \end{aligned} \tag{2-3}$$

$$\begin{aligned}
 \log \left[\frac{p(Y = \text{삼태아이상})}{P(Y = \text{단태아})} \right] &= \beta_0 + \beta_1 \text{부연령}_i + \beta_2 \text{모연령}_i + \beta_3 \text{임신주수}_i \\
 &+ \beta_4 \text{동거기간}_i + \beta_5 \text{다태아출산순위}_i + \beta_6 \text{모충출생아수}_i \\
 &+ \beta_7 I_{11}(\text{출생장소} = \text{주택}) + \beta_8 I_{12}(\text{출생장소} = \text{병원}) \\
 &+ \beta_9 I_{31}(\text{부교육} = \text{무학}) + \beta_{10} I_{32}(\text{부교육} = \text{초졸}) + \beta_{11} I_{33}(\text{부교육} = \text{중졸}) \\
 &+ \beta_{12} I_{34}(\text{부교육} = \text{고졸}) + \beta_{13} I_{35}(\text{부교육} = \text{대학}) \\
 &+ \beta_{14} I_{41}(\text{모교육} = \text{무학}) + \beta_{15} I_{42}(\text{모교육} = \text{초졸}) + \beta_{16} I_{43}(\text{모교육} = \text{중졸}) \\
 &+ \beta_{17} I_{44}(\text{모교육} = \text{고졸}) + \beta_{18} I_{45}(\text{모교육} = \text{대학}) + e_i
 \end{aligned} \tag{2-4}$$

$$\begin{aligned}
 \log \left[\frac{p(Y = \text{쌍태아})}{P(Y = \text{단태아})} \right] &= \beta_0 + \beta_1 \text{부연령}_i + \beta_2 \text{모연령}_i + \beta_3 \text{임신주수}_i \\
 &+ \beta_4 \text{동거기간}_i + \beta_5 \text{다태아출산순위}_i + \beta_6 \text{모충출생아수}_i \\
 &+ \beta_7 I_{11}(\text{출생장소} = \text{주택}) + \beta_8 I_{12}(\text{출생장소} = \text{병원}) \\
 &+ \beta_9 I_{31}(\text{부교육} = \text{무학}) + \beta_{10} I_{32}(\text{부교육} = \text{초졸}) + \beta_{11} I_{33}(\text{부교육} = \text{중졸}) \\
 &+ \beta_{12} I_{34}(\text{부교육} = \text{고졸}) + \beta_{13} I_{35}(\text{부교육} = \text{대학}) \\
 &+ \beta_{14} I_{41}(\text{모교육} = \text{무학}) + \beta_{15} I_{42}(\text{모교육} = \text{초졸}) + \beta_{16} I_{43}(\text{모교육} = \text{중졸}) \\
 &+ \beta_{17} I_{44}(\text{모교육} = \text{고졸}) + \beta_{18} I_{45}(\text{모교육} = \text{대학}) + e
 \end{aligned} \tag{2-5}$$

주 : 동거기간, 부·모연령, 임신주수는 정규성을 위해 변수변환(log-transformation)을 함

2. 출생방정식 연도별 비교

가. 모총출생아수 출생방정식 결과와 비교

- 6개년 모총출생아수의 출생방정식 도출결과 동거기간, 다태아출산순위, 모교육정도, 다태아여부(단태아), 출생아체중, 모연령, 부교육정도는 모총출생아수에 양(+)의 영향을 임신주수, 다태아여부(쌍태아)는 음(-)의 영향을 보임(<표 2-2>, <표 2-3> 참조)

<표 2-2> '07~ '12년 연도별 회귀분석 결과

2007년도($F=52696.52^{***}$, $R^2=0.46$)

$$\begin{aligned} \text{모총출생아수}_i &= 0.598\log\text{동거기간}_i + 0.169\text{다태아출산순위}_i + 0.076\text{모교육(고졸)}_i \\ &+ 0.104\text{다태아여부(단태아)}_i - 0.082\log\text{임신주수}_i + 0.06\text{출생아체중}_i \\ &+ 0.071\log\text{부연령}_i + 0.049\text{부교육(고졸)}_i \end{aligned}$$

2008년도($F=34909.39^{***}$, $R^2=0.41$)

$$\begin{aligned} \text{모총출생아수}_i &= 0.603\log\text{동거기간}_i + 0.164\text{다태아출산순위}_i + 0.078\text{모교육(고졸)}_i \\ &+ 0.097\text{다태아여부(단태아)}_i + 0.071\log\text{모연령}_i - 0.089\log\text{임신주수}_i \\ &+ 0.063\text{출생아체중}_i + 0.0497\text{부교육(고졸)}_i \end{aligned}$$

2009년도($F=44717.22^{***}$, $R^2=0.45$)

$$\begin{aligned} \text{모총출생아수}_i &= 0.601\log\text{동거기간}_i + 0.155\text{다태아출산순위}_i + 0.076\text{모교육(고졸)}_i \\ &- 0.084\text{다태아여부(쌍태아)}_i + 0.064\log\text{모연령}_i - 0.093\log\text{임신주수}_i \\ &+ 0.066\text{출생아체중}_i + 0.05\text{부교육(고졸)}_i \end{aligned}$$

2010년도($F=35282.04^{***}$, $R^2=0.41$)

$$\begin{aligned} \text{모총출생아수}_i &= 0.599\log\text{동거기간}_i + 0.157\text{다태아출산순위}_i + 0.08\text{모교육(고졸)}_i \\ &+ 0.06\log\text{모연령}_i + 0.083\text{다태아여부(단태아)}_i - 0.097\log\text{임신주수}_i \\ &+ 0.072\text{출생아체중}_i + 0.052\text{부교육(고졸)}_i \end{aligned}$$

2011년도($F=38077.09^{***}$, $R^2=0.42$)

$$\begin{aligned} \text{모총출생아수}_i &= 0.617\log\text{동거기간}_i + 0.16\text{다태아출산순위}_i + 0.077\text{모교육(고졸)}_i \\ &- 0.091\text{다태아여부(쌍태아)}_i + 0.043\log\text{모연령}_i - 0.097\log\text{임신주수}_i \\ &+ 0.071\text{출생아체중}_i + 0.05\text{부교육(고졸)}_i \end{aligned}$$

2012년도($F=33006.90^{***}$, $R^2=0.38$)

$$\begin{aligned} \text{모총출생아수}_i &= 0.609\log\text{동거기간}_i + 0.174\text{다태아출산순위}_i + 0.075\text{모교육(고졸)}_i \\ &+ 0.101\text{다태아여부(단태아)}_i + 0.039\log\text{모연령}_i - 0.103\log\text{임신주수}_i \\ &+ 0.075\text{출생아체중}_i + 0.05\text{부교육(고졸)}_i \end{aligned}$$



- 특히 동거기간과 다태아출산순위는 모총출생아수에 대해 다른 변수들의 영향보다 약 6배, 1.7배 높게 분석됨
 - 모교육정도와 단태아여부는 약 0.7배, 1.0배 정도 나타남
- 다태아여부의 경우 2009년, 2011년에는 타연도와 상이하게 쌍태아가 모총출생아수에 음(-)의 영향을 미치는 것으로 분석됨
- 2007년에는 다른 5개년과 다르게 모연령이 아닌 부연령이 모총출생아수에 영향을 미치는 변수로 도출됨

〈표 2-3〉 연도별 회귀분석 결과 변수 우선 순위[음의영향 표기→ ‘(-)’]

2007년	① 동거기간, ② 다태아출산순위, ③ 모교육정도(고졸), ④ 다태아여부(단태아), ⑤ 임신주수(-), ⑥ 출생아체중, ⑦ 부연령, ⑧ 부교육정도(고졸)
2008년	① 동거기간, ② 다태아출산순위, ③ 모교육정도(고졸) ④ 다태아여부(단태아), ⑤ 모연령, ⑥ 임신주수(-), ⑦ 출생아체중, ⑧ 부교육정도(고졸)
2009년	① 동거기간, ② 다태아출산순위, ③ 모교육정도(고졸), ④ 다태아여부(쌍태아)(-), ⑤ 모연령, ⑥ 임신주수(-), ⑦ 출생아체중, ⑧ 부교육정도(고졸)
2010년	① 동거기간, ② 다태아출산순위, ③ 모교육정도(고졸), ④ 모연령, ⑤ 다태아여부(단태아), ⑥ 임신주수(-), ⑦ 출생아체중, ⑧ 부교육정도(고졸)
2011년	① 동거기간, ② 다태아출산순위, ③ 모교육정도(고졸), ④ 다태아여부(쌍태아)(-), ⑤ 모연령, ⑥ 임신주수(-), ⑦ 출생아체중, ⑧ 부교육정도(고졸)
2012년	① 동거기간, ② 다태아출산순위, ③ 모교육정도(고졸), ④ 다태아여부(단태아), ⑤ 모연령, ⑥ 임신주수(-), ⑦ 출생아체중, ⑧ 부교육정도(고졸)

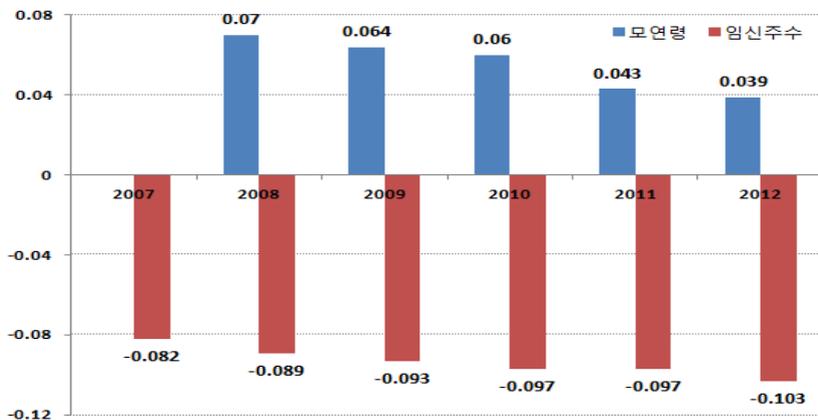
주: ***(p -value<0.001), **(p -value<0.01), *(p -value<0.05),

자세한 회귀분석 결과는 [부록]에 제시

- 앞절에서 도출된 모총출생아수에 대한 주요변수들의 관계를 살펴보면 다음과 같음

- 다른 변수들의 영향력이 일정하다고 가정할 때, 동거기간이 1년 증가할 때 모출생아수는 평균 0.6명 정도 상승
- 다태아출산순위가 첫째→둘째, 둘째→셋째로 변화될 때 모총출생아수는 0.17명 정도 상승
- 모연령 변수의 표준화 회귀계수값이 점진적 하향을 보여 영향력이 점차 감소하는 것을 알 수 있으며, 2008년에는 0.07, 2009년 0.064, 2010년 0.06, 2011년 0.043, 2012년 0.039로 연평균 -13.6%의 하강기조를 보임
 - 이는 모연령이 증가할수록 모총출생아수는 낮아지는 일반적 견해를 지지함
- 임신주수는 모연령과 동일한 패턴을 보이는 것으로 나타났으며, 2007년 -0.082, 2008년 -0.089, 2009년 -0.093, 2010년 -0.097, 2011년 -0.097, 2012년 -0.103으로 연평균 -5.9%를 보임
- 교육정도를 살펴보면 부·모교육정도가 고졸인 경우 다른 학력에 비해 상대적으로 높은 모총출생아수를 나타내는 결과가 도출
 - 부교육정도보다는 모교육정도가 모총출생아수에 미치는 영향력이 크다고 할 수 있으며, 많게는 1.6배 적게는 1.3배 정도로 나타남

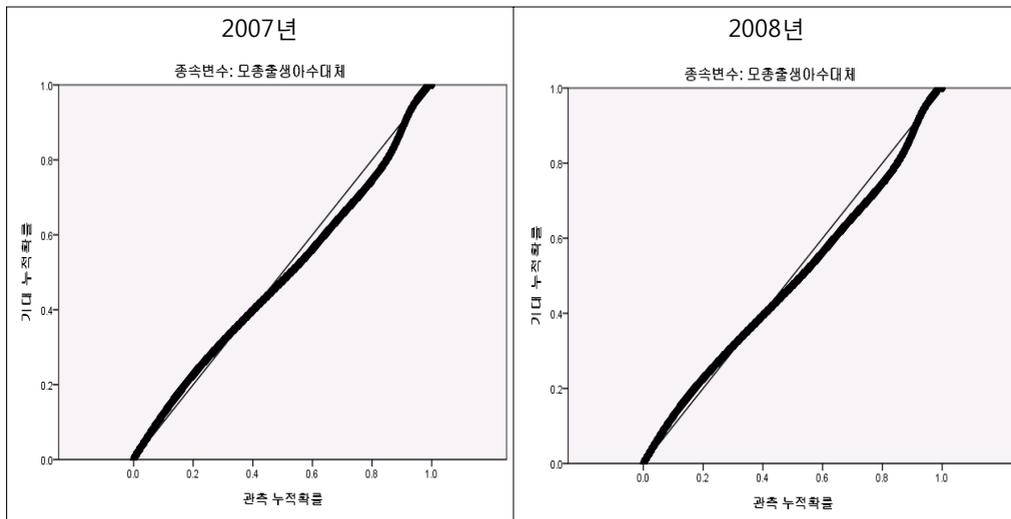
[그림 2-2] 모총출생아수에 대한 모연령과 임신주수 영향력 추이





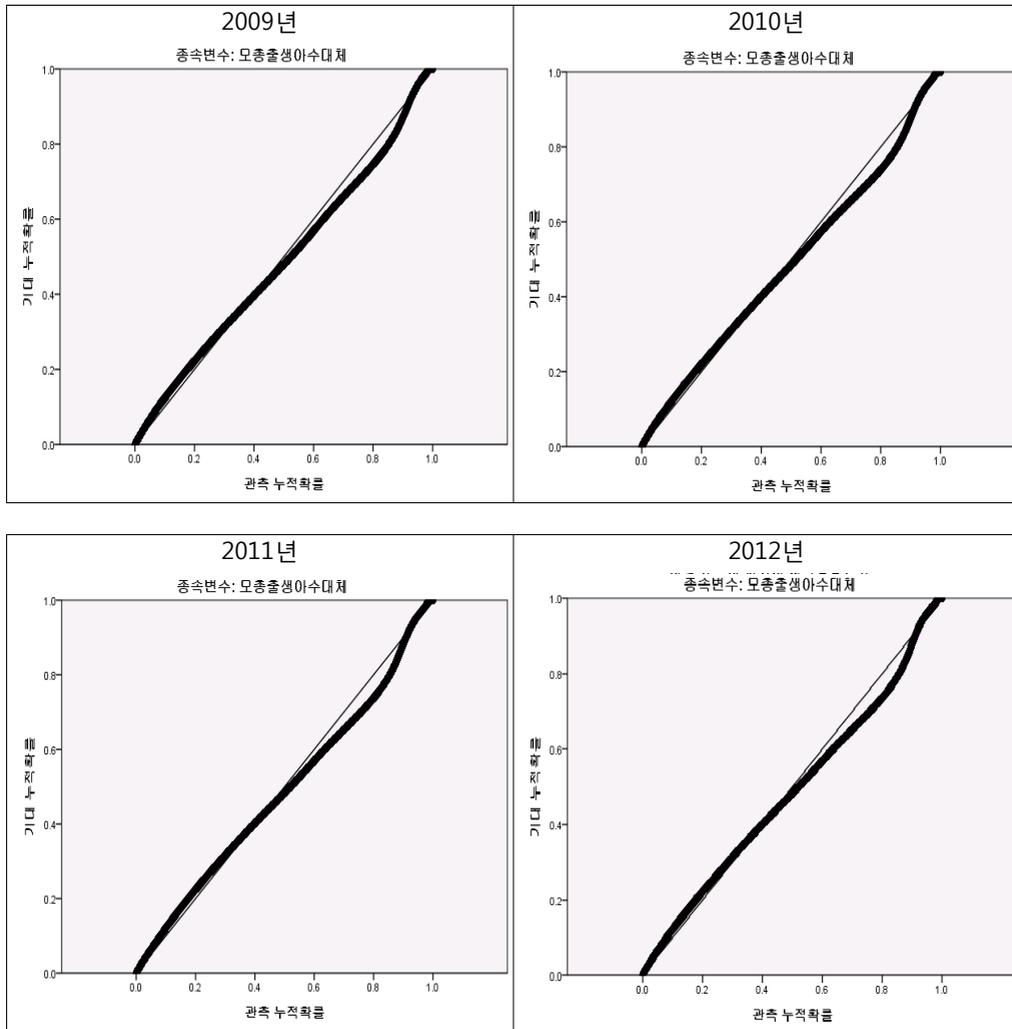
- 6개년 분석결과를 토대로 개괄적인 모총출생아수의 출생방정식을 제안해 본다면, 동거기간, 다태아출산순위, 부·모교육정도(고졸), 다태아여부(단태아, 쌍태아), 모연령, 임신주수, 출생아체중의 변수로 구성됨
 - 모총출생아수에 대한 주요 변수들의 영향력은 동거기간이 약 0.6, 다태아출산순위 약 0.17, 모교육정도는 약 0.075, 다태아여부는 단태아(쌍태아) 약 0.9(-0.9), 모연령은 약 0.04, 임신주수는 약 -0.10 이하, 출생아체중은 약 0.075, 부교육정도는 약 0.05일 것으로 추정할 수 있음
- 도출된 모총출생아수 출생방정식의 정규성을 알아보기 위해 적합회귀 표준화 잔차의 정규 p-p 도표⁴⁾를 이용하여 살펴보면, 정규성에 크게 위배되지 않는다는 것을 알 수 있음

[그림 2-3] 모총출생아수 출생방정식의 p-p plot



4) p-p 도표(또는 q-q plot)는 관측치의 분포가 정규성을 만족하는지의 여부를 알아보기 위한 것으로, 45도 직선에 크게 벗어나지 않으면 정규성에 위배되지 않는다는 것을 의미한다. 여기서는 적합 후 잔차의 정규성을 살펴보았는데 직선에 크게 벗어나지 않는 것으로 판단되었다([그림 2-3] 참조).

[그림 2-3] 모종출생아수 출생방정식의 p-p plot(계속)



나. 다태아여부 출생방정식 결과와 비교

- 다태아여부의 다항로지스틱 회귀모형 적합결과를 살펴보면 다음과 같음
 - 2007년에 다태아출산순위, 동거기간, 모연령, 부연령, 출생장소(병원)은 단태아인 경우가 쌍태아, 삼태아 이상 보다 순서적으로 높은 확률을 가진다고 할 수 있음



- 즉, 단태아인 경우 다태아출산순위, 동거기간, 모연령, 부연령이 쌍태아, 삼태아 이상인 경우보다 높은 값으로 나타나는 경향이 있으며 자택이나 기타장소에서 태어나는 것보다 병원에서 태어나는 경우가 많음
 - 2008년은 다태아출산순위, 동거기간, 모연령, 출생장소(병원)가 단태아인 경우 쌍태아, 삼태아 이상보다 순서적으로 높은 확률을 가짐
 - 2009년은 다태아출산순위, 동거기간, 모연령, 부연령이 단태아인 경우 쌍태아, 삼태아 이상보다 순서적으로 높은 확률을 가진 반면, 2010년은 다태아출산순위, 동거기간, 모연령, 임신주수가 단태아인 경우 쌍태아, 삼태아 이상보다 순서적으로 높은 확률을 나타냄
 - 2011년은 다태아출산순위, 동거기간, 부연령, 임신주수가 단태아인 경우 쌍태아, 삼태아 이상보다 순서적으로 높은 확률을 가진 반면, 2012년은 임신주수, 동거기간, 부연령, 모충출생아수가 단태아인 경우 쌍태아, 삼태아 이상보다 순서적으로 높은 확률을 보임
- 특이한 점은 2007~2011년까지 다태아출산순위, 동거기간, 부·모연령이 순서적으로 높은 확률을 가지는 변수로 나타났으며, 5개년 동안 가장 큰 영향요인으로 도출된 다태아출산순위는 2012년에는 통계적으로 유의하지 않게 나타남
- 이것은 다른 요인에 기인한 것으로 판단되며 부가적으로 가족력, 유전 정보 등의 추가정보를 반영해 본다면 보다 정밀한 함수 관계식을 유도할 수 있을 것임

〈표 2-4〉 연도별 다태아여부 출생방정식 결과

2007년 적합 ($\chi^2=73650^{***}$, 최종모형:-2log우도=52620)					2008년 적합 ($\chi^2=68220^{***}$, 최종모형:-2log우도=50820)				
다태아여부	B	유의확률	Exp(B)	다태아여부	B	유의확률	Exp(B)		
쌍 태 아				쌍 태 아					
절편	-8.233	.000		절편	-6.351	.000			
출생장소	.464	.000	1.59	출생장소	.493	.000	1.63		
모총출생아수	-1.484	.000	.22	모총출생아수	-1.414	.000	.24		
다태아출산순위	9.810	.000	18206.83	다태아출산순위	10.486	.000	35827.51		
log모연령	1.252	.000	3.49	log모연령	1.431	.000	4.18		
부교육정도	.112	.000	1.11	모교육정도	.102	.000	1.10		
log부연령	1.117	.001	3.05	부교육정도	.001	.950	1.00		
log임신주수	-1.430	.014	.23	log임신주수	-2.364	.000	.09		
출생아체중	-2.514	.000	.08	출생아체중	-2.417	.000	.08		
log동거기간	2.166	.000	8.72	log동거기간	2.118	.000	8.31		
삼 태 아 이 상				삼 태 아 이 상					
절편	-21.859	.000		절편	3.155	.543			
출생장소	-.198	.764	.82	출생장소	.048	.938	1.04		
모총출생아수	-1.319	.000	.26	모총출생아수	-2.280	.000	.10		
다태아출산순위	11.997	.000	162324.92	다태아출산순위	13.779	.000	964014.83		
log모연령	.595	.768	1.81	log모연령	-2.165	.227	.11		
부교육정도	.181	.218	1.19	모교육정도	.495	.000	1.64		
log부연령	8.324	.000	4120.59	부교육정도	-.329	.007	.72		
log임신주수	-2.182	.445	.11	log임신주수	-9.272	.003	0.0		
출생아체중	-4.111	.000	.01	출생아체중	-3.149	.000	.04		
log동거기간	2.926	.000	18.65	log동거기간	2.631	.000	13.88		
참조 범주는 단태아				참조 범주는 단태아					

2009년 직합 ($\chi^2=59830^{***}$, 최종 모형:-2log우도=52660)					2010년 직합 ($\chi^2=67910^{***}$, 최종 모형:-2log우도=52010)										
다태아여부	B	유의확률	Exp(B)	다태아여부	B	유의확률	Exp(B)								
쌍태아				쌍태아				절편	-10.706	.000					
								모총출생아수	-1.379	.000	.25	출생장소	.281	.003	1.32
								다태아출산순위	11.482	.000	96927.45	모총출생아수	-1.092	.000	.33
								log모연령	1.144	.001	3.13	다태아출산순위	11.332	.000	83440.60
								모교육정도	.067	.001	1.06	log모연령	.910	.001	2.48
								log부연령	.818	.012	2.26	모교육정도	.076	.000	1.07
								log임신주수	-34.795	.000	0.0	log임신주수	.738	.208	2.09
								log동거기간	1.891	.000	6.62	출생아체중	-2.549	.000	.07
								절편	65.186	.000		log동거기간	1.806	.000	6.08
								모총출생아수	-456	.000	.63	절편	-16.787	.000	
삼태아이상				삼태아이상				출생장소	.029	.959	1.02				
								다태아출산순위	13.215	.000	548569.66	모총출생아수	-.855	.000	.42
								log모연령	-1.715	.459	.18	다태아출산순위	13.428	.000	678855.20
								모교육정도	.344	.006	1.41	log모연령	1.916	.235	6.79
								log부연령	-.900	.688	.40	모교육정도	-.115	.308	.89
								log임신주수	-55.908	.000	0.0	log임신주수	1.419	.577	4.13
								log동거기간	2.515	.000	12.36	출생아체중	-4.286	.000	.01
								절편				log동거기간	2.087	.000	8.06
								모총출생아수				참조 범주는 단태아			
								참조 범주는 단태아				참조 범주는 단태아			

2011년 직합
($\chi^2=68050^{***}$, 최종모형:-2log우도=52070)

다태아여부	B	유의확률	Exp(B)
쌍 태 아			
절편	-10.082	.000	
모총출생아수	-1.227	.000	.29
다태아출산순위	11.536	.000	102352.54
모교육정도	.151	.000	1.16
log부연령	.920	.001	2.51
log임신주수	.356	.538	1.42
출생아체중	-2.539	.000	.07
log동거기간	2.064	.000	7.87
삼 태 아 이 상			
절편	-10.433	.008	
모총출생아수	-.252	.005	.77
다태아출산순위	12.924	.000	410086.38
모교육정도	.015	.886	1.01
log부연령	-6.183	.000	.00
log임신주수	4.960	.035	142.56
출생아체중	-4.554	.000	.01
log동거기간	2.674	.000	14.49

참조 범주는 단태아

2012년 직합
($\chi^2=41450^{***}$, 최종모형:-2log우도=98730)

다태아여부	B	유의확률	Exp(B)
쌍 태 아			
절편	-4.219	.000	
출생장소	.191	.010	1.21
모총출생아수	.350	.000	1.41
log모연령	.331	.140	1.39
모교육정도	.273	.000	1.31
부교육정도	.051	.000	1.05
log부연령	1.078	.000	2.93
log임신주수	3.089	.000	21.95
출생아체중	-3.107	.000	.04
log동거기간	.529	.000	1.69
삼 태 아 이 상			
절편	-4.454	.204	
출생장소	.713	.185	2.04
모총출생아수	.930	.000	2.53
log모연령	-6.428	.000	.00
모교육정도	.382	.000	1.46
부교육정도	-.043	.561	.95
log부연령	.346	.787	1.41
log임신주수	8.453	.000	4690.54
출생아체중	-4.582	.000	.01
log동거기간	.857	.000	2.35

참조 범주는 단태아

주: ***(p -value<0.001), **(p -value<0.01), *(p -value<0.05)



Ⅲ. 결론 및 활용성 제언

- 본 연구는 출산에 관련된 차별출산력의 선행연구와는 다른 출생신고자료 항목들 간의 연도별 관계규명을 도출하여 출생방정식을 제안
 - 흔히 소득, 교육정도, 연령이 출산력에 영향변수로 규명되는 차별출산력 연구는 단편적인 변수 규명에 초점이 맞춰져 있으나, 본 연구는 출생신고자료를 활용하여 주요항목들 간의 함수관계를 도출한 것에 의미가 있음

- 현재 인구통계는 항목별 단순 수치와 분포만을 제시하고 있어 이해 측면에서는 쉽다는 장점이 있으나 주요항목의 개인별 특성치를 반영하는 데는 한계점이 있으므로 본 연구에서는 출생방정식의 활용성을 제언함
 - 출생방정식은 출생신고자료 주요항목의 개인별 특성치를 반영할 수 있어 개인 맞춤형 출생방정식으로도 적용이 가능함⁵⁾
 - 예를 들어 출생방정식에 개인별 특성치(부·모연령, 동거기간, 다태아출산순위, 모교육, 다태아여부, 임신주수, 출생아체중, 부연령, 부교육)를 반영하면 모총출생아수 및 다태아여부의 예측값이 도출되어 가족 구성원과 다태아여부 형태를 파악할 수 있음
 - 다양한 영향요인 값(가령 학력, 연령, 임신주수, 출생아체중 등의 다른 임의값)을 출생방정식에 반영한다면 모총출생아수와 다태아여부의 확률값 변화를 보여줄 수 있으므로 시나리오 분석이 가능
 - 첫째 아이를 가진 부모가 둘째를 계획하고 있는 경우 모총출생아수 예측값이 2 근방으로 도출되는 연도별 영향요인 값을 살펴본다면 출산 계획에 참고자료로 활용이 가능
 - 쌍태아를 계획하고 있는 경우 다태아여부의 출생방정식을 활용해 영향요인 값을 참고할 수 있음

5) 다양한 개인별 시나리오 분석 결과는 [부록]에 제시

- 더 나아가 보험업계의 활용을 제안해 본다면 어린이보험(태아보험 포함) 가입 유도 시 보험설계자가 가입자의 출생신고자료의 특성값을 출생방정식에 반영해 현재의 가족구성원과 향후의 출산계획을 사전에 파악할 수 있으므로 보험가입동기를 이끌 수 있을 것임

□ 끝으로 본 논문의 한계 및 향후과제에 대해 제시하면 아래와 같음

- 출생신고자료의 주요항목들 중 모총출생아수, 다태아여부의 출생방정식을 제안하였으나 보다 체계적이고 종합적인 출생 규명을 위해 9개 주요항목별 출생방정식 도출이 필요
- 제시한 출생방정식은 출생신고자료만을 이용하여 산출하였으므로 좀 더 대표성이 담보된 자료(소득, 가족력 정보)를 반영할 수 있다면 보다 정밀한 출생방정식을 도출할 수 있을 것임
- 출생신고서의 기입 항목이 시간의 흐름에 따라 점차 확대되지만 주요변수들 간의 시계열적인 일관성과 용어 통일성이 보장된다면 집단 코호트별(1980, 1990, 2000, 2010년대 출생집단별) 출생방정식을 도출하여 비교하는 것도 유용할 것이며, 이들에 대한 연구는 향후 과제로 남기고자 함



【참고문헌】

- 민희철(2008), “임금 및 소득이 출산 간격에 미치는 효과의 분석”, 재정학연구1(1), pp.41~61.
- 신혜원·최명섭·김의준(2009), “서울시 추가자녀 출산계획에 미치는 가구 특성의 요인분석”, 서울도시연구, 10(2), pp.33~47.
- 최은영·박영실(2010), “사회경제적 특성과 지역별 차별 출산력 분석”, 통계개발원 연구보고서.
- 최은영, 박영실(2011), “소득과 자산에 따른 차별 출산력”, 통계개발원 연구보고서.
- 통계청(2011), 「2011년 인구동태통계연보」, 통계청.
- 통계청(2012), 「2012년 인구동태통계연보」, 통계청.
- 통계청(2012), 「2010~2060 장래인구추계」, 통계청.
- 통계청(2013), 「2012년 생명표」, 통계청.
- 통계청(2013), 「인구통계 기초와 응용」, 통계청.
- Bongaarts(1998), “Fertility and Reproductive Preferences in Post-Transition Societies.”, Population Council Working paper No. 114. A paper read for the Conference on Global Fertility Transition, Bellagio, Italy, may, 1998.
- Robinson(1997), “The economic theory of fertility over three decades”, Population Studies, 51, pp.63~74.
- Sleebos, Joellw(2003), “Low Fertility Rates in OECD Countries: Facts and Policy Response”, OECD Social, Employment and Migration Working Paper.

【부 록】

〈부표 2-1〉 출생신고자료 각 연도별 기초통계

연도	항목	부연령 (단위:세)	모연령 (단위:세)	임신주수 (단위:주)	다태아 출산순위	모총 출생아수	출생아체중 (단위:kg)
2007		491,951	493,584	492,260	494,280	493,378	494,032
		33.42(4.37)	30.61(3.99)	38.91(1.60)	1.01(0.12)	1.57(0.69)	3.23(0.45)
2008		460,995	462,994	463,799	463,743	462,436	463,550
		33.64(4.41)	30.82(4.01)	38.80(1.60)	1.01(0.12)	1.59(0.70)	3.22(0.45)
2009		440,096	442,318	442,809	443,044	442,136	442,861
		33.81(4.44)	31.00(4.01)	38.77(1.61)	1.01(0.12)	1.59(0.70)	3.22(0.45)
2010		464,853	467,422	467,927	468,150	467,136	467,962
		34.02(4.48)	31.28(4.04)	38.73(1.62)	1.01(0.12)	1.62(0.72)	3.22(0.47)
2011		466,505	469,890	469,961	470,178	469,608	470,000
		34.16(4.56)	31.44(4.10)	38.69(1.62)	1.01(0.12)	1.62(0.73)	3.21(0.46)
2012		479,253	483,436	483,267	483,474	483,006	483,339
		34.30(4.54)	31.62(4.09)	38.65(1.63)	1.02(0.13)	1.61(0.72)	3.21(0.46)

〈부표 2-1〉 출생신고자료 각 연도별 기초통계(계속)

연도	항목	부교육정도 ¹⁾	모교육정도	동거기간 (단위:년)	출생장소	다태아여부
2007		491,322	492,525	488,680	494,419	494,280
		고졸이하(35.6%) 대졸이상(64.4%)	고졸이하(39.3%) 대졸이상(60.7%)	3.41(3.06)	자택(1.1%), 병원(98.6%), 기타(0.3%)	단태아(97.2%), 쌍태아(2.7%), 삼태아이상(0.1%)
2008		460,469	461,854	463,799	463,550	463,743
		고졸이하(33.2%) 대졸이상(66.8%)	고졸이하(36.5%) 대졸이상(63.5%)	3.44(3.42)	자택(1.3%), 병원(98.2%), 기타(0.5%)	단태아(97.2%), 쌍태아(2.7%), 삼태아이상(0.1%)
2009		439,146	441,036	443,068	442,829	443,045
		고졸이하(31.6%) 대졸이상(68.4%)	고졸이하(34.3%) 대졸이상(65.7%)	3.41(3.08)	자택(1.3%), 병원(98.2%), 기타(0.5%)	단태아(97.2%), 쌍태아(2.7%), 삼태아이상(0.1%)
2010		464,017	466,249	468,182	467,972	468,150
		고졸이하(30.3%) 대졸이상(69.7%)	고졸이하(32.6%) 대졸이상(67.4%)	3.53(3.45)	자택(1.3%), 병원(98.2%), 기타(0.5%)	단태아(97.2%), 쌍태아(2.7%), 삼태아이상(0.1%)
2011		465,988	469,178	470,187	469,969	470,178
		고졸이하(29.3%) 대졸이상(70.7%)	고졸이하(31.0%) 대졸이상(69.0%)	3.52(3.41)	자택(1.3%), 병원(98.2%), 기타(0.5%)	단태아(97.0%), 쌍태아(2.8%), 삼태아이상(0.2%)
2012		478,696	482,698	483,483	483,292	483,474
		고졸이하(27.7%) 대졸이상(72.3%)	고졸이하(28.9%) 대졸이상(71.1%)	3.46(3.57)	자택(1.1%), 병원(98.5%), 기타(0.4%)	단태아(96.8%), 쌍태아(3.1%), 삼태아이상(0.1%)

주 : 변수별 첫 라인은 출생신고자료 항목별 건수, 둘째 라인은 평균과 표준편차 및 비중임.

다태아 출산순위의 단태아(1), 쌍태아(2), 삼태아이상(3), 모종출생아수는 1아, 2아, ..., 8아 이상으로 표기

1) 교육정도의 원자료는 초졸-대졸, 대학원이상이며, 지면관계상 고졸이하, 대졸이상으로 구분

<부표 2-2> 출생신고자료에 대한 일반선형모형과 변수선택 회귀분석 결과 비교

2007년		일반선형모형		회귀분석		2008년		일반선형모형		회귀분석		2009년		일반선형모형		회귀분석	
Var.		계수(se)		계수(표준화계수)		Var.		계수(se)		계수(표준화계수)		Var.		계수(se)		계수(표준화계수)	
log동거기간		0.996(0.002)		0.996(0.598)		log동거기간		1.017(0.002)		1.017(0.603)		log동거기간		1.028(0.002)		1.028(0.601)	
다태아출산순위		0.992(0.009)		0.976(0.169)		다태아출산순위		0.979(0.009)		0.974(0.164)		다태아출산순위		0.924(0.009)		0.925(0.155)	
모교육		-0.110(0.002)		0.110(0.076)		모교육		-0.115(0.002)		0.115(0.078)		모교육		-0.114(0.002)		0.114(0.076)	
다태아여부		-0.455(0.006)		0.444(0.104)		다태아여부		-0.420(0.007)		0.416(0.097)		다태아여부		0.366(0.007)		-0.368(-0.084)	
log부연령		0.877(0.014)		0.877(0.014)		log모연령		0.861(0.015)		0.861(0.071)		log모연령		0.784(0.015)		0.784(0.064)	
log임신주수		-3.047(0.047)		-3.041(-0.082)		log임신주수		-3.290(0.049)		-3.288(-0.089)		log임신주수		-3.421(0.05)		-3.421(-0.093)	
출생이체중		0.091(0.002)		0.092(0.060)		출생이체중		0.097(0.002)		0.097(0.063)		출생이체중		0.102(0.002)		0.102(0.066)	
부교육		-0.073(0.002)		0.072(0.049)		부교육		-0.074(0.002)		0.074(0.049)		부교육		-0.077(0.002)		0.077(0.050)	
상수		3.552(0.076)		2.932()		상수		3.997(0.079)		3.392()		상수		3.992(0.081)		4.166()	
F-value		50343.9***		52696.5***		F-value		49142.4***		34909.4***		F-value		45605.0***		44717.2***	
R ²		0.449		0.460		R ²		0.459		0.410		R ²		0.452		0.450	
2010년		일반선형모형		회귀분석		2011년		일반선형모형		회귀분석		2012년		일반선형모형		회귀분석	
Var.		계수(se)		계수(표준화계수)		Var.		계수(se)		계수(표준화계수)		Var.		계수(se)		계수(표준화계수)	
log동거기간		1.025(0.002)		1.025(0.599)		log동거기간		1.051(0.002)		1.051(0.617)		log동거기간		1.033(0.002)		1.033(0.609)	
다태아출산순위		0.958(0.009)		0.956(0.157)		다태아출산순위		0.950(0.009)		0.950(0.160)		다태아출산순위		0.977(0.008)		0.976(0.174)	
모교육		-0.125(0.002)		0.126(0.08)		모교육		-0.123(0.002)		0.123(0.077)		모교육		-0.123(0.002)		0.123(0.075)	
다태아여부		-0.367(0.007)		0.365(0.083)		다태아여부		0.396(0.007)		-0.396(-0.091)		다태아여부		-0.415(0.006)		0.414(0.101)	
log모연령		0.760(0.015)		0.759(0.06)		log모연령		0.539(0.015)		0.539(0.043)		log모연령		0.483(0.015)		0.483(0.039)	
log임신주수		-3.642(0.05)		-3.641(-0.097)		log임신주수		-3.632(0.05)		-3.632(-0.097)		log임신주수		-3.801(0.049)		-3.801(-0.103)	
출생이체중		0.114(0.002)		0.114(0.072)		출생이체중		0.114(0.002)		0.114(0.071)		출생이체중		0.119(0.002)		0.119(0.075)	
부교육		-0.083(0.002)		0.083(0.052)		부교육		-0.082(0.002)		0.082(0.05)		부교육		-0.083(0.002)		0.083(0.050)	
상수		4.705(0.081)		4.131()		상수		4.622(0.08)		4.813()		상수		5.331(0.079)		4.712()	
F-value		48021.1***		35282.04***		F-value		50546.1***		38077.1***		F-value		49435.1***		33006.9***	
R ²		0.451		0.410		R ²		0.462		0.420		R ²		0.450		0.460	

주: ***(p-value<0.001), **(p-value<0.01), *(p-value<0.05)

일반선형모형(GLM)과 Stepwise 방법을 적용한 회귀분석의 결과이며, 두 분석방법을 비교한 결과 매우 유사한 것으로 도출된다. 일반선형모형의 모교육, 부교육, 다태아여부 변수의 기준범주는 0이므로 적합결과 부호가 반대로 도출(1에 비해 상대적으로 음의 영향을 나타냄)

〈부표 2-3〉 출생방정식을 활용한 다양한 개인별 시나리오 결과

No.	시나리오
1	<p>2010년 35세 부연령과 32세 모연령이 결혼했을 경우(동거기간 1년, 부모교육정도 고졸) 태어난 모총출생아수는 얼마일까?</p> <p>가정 : 타변수들의 값은 35세 부연령과 32세 모연령의 평균값을 입력 임신주수 : 38.71주, 출생아체중 : 3.23kg, 다태아출산순위=다태아여부=1</p> <p>→ 첫아이가 태어날 확률은 약 55% 정도임</p> <p>→ 2010년도 모총출생아수 출생방정식을 활용하여 계산한 결과</p> $\begin{aligned} \text{모총출생아수}_i &= 0.5991\log(\text{동거기간}_i + 0.157\text{다태아출산순위}_i + 0.08\text{모교육(고졸)}_i) \\ &+ 0.061\log(\text{모연령}_i + 0.083\text{다태아여부(단태아)}_i) - 0.0971\log(\text{임신주수}_i \\ &+ 0.072\text{출생아체중}_i + 0.052\text{부모교육(고졸)}_i) \end{aligned}$ $0.5404 = 0.5991\log(1) + 0.157(1) + 0.08\text{모교육} + 0.061\log(32) + 0.083(1) - 0.0971\log(38.71) + 0.072(3.23) + 0.052\text{부모교육}$
2	<p>2010년 35세 부연령과 32세 모연령이 결혼했을 경우(동거기간 3년, 부모교육정도 대졸) 태어난 모총출생아수는 얼마일까?</p> <p>가정 : 타변수들의 값은 35세 부연령과 32세 모연령의 평균값을 입력 임신주수 : 38.71주, 출생아체중 : 3.23kg, 다태아출산순위=다태아여부=1</p> <p>→ 첫아이가 태어날 확률은 약 70% 정도임</p> <p>→ 2010년도 모총출생아수 출생방정식을 활용하여 계산한 결과</p> $\begin{aligned} \text{모총출생아수}_i &= 0.5991\log(\text{동거기간}_i + 0.157\text{다태아출산순위}_i + 0.08\text{모교육(고졸)}_i) \\ &+ 0.061\log(\text{모연령}_i + 0.083\text{다태아여부(단태아)}_i) - 0.0971\log(\text{임신주수}_i \\ &+ 0.072\text{출생아체중}_i + 0.052\text{부모교육(고졸)}_i) \end{aligned}$ $0.694 = 0.5991\log(3) + 0.157(1) + 0.061\log(32) + 0.083(1) - 0.0971\log(38.71) + 0.072(3.23)$

3	<p>2007년 27세 부연령과 30세 모연령이 결혼했을 경우(연상연하커플, 동거기간 2년, 부모교육정도 대졸) 태어날 모총출생아수는? 가정 : 타변수들의 값은 27세 부연령과 30세 모연령의 평균값을 입력 → 첫아이가 태어날 확률은 약 62% 정도임 → 2007년도 모총출생아수 출생방정식을 활용하여 계산한 결과 모총출생아수_i = 0.598log동거기간_i + 0.169다태아출산순위_i + 0.076모교육(고졸)_i + 0.104다태아여부(단태아)_i - 0.082log임신주수_i + 0.06출생아체중_i + 0.071log부연령_i + 0.049부모교육(고졸)_i 0.618 = 0.598log(2) + 0.169(1) + 0.104(1) - 0.082log(38.91) + 0.06(3.24) + 0.071log(27)</p>
4	<p>2007년 27세 부연령과 30세 모연령이 결혼했을 경우(연상연하커플, 동거기간 3년, 부모교육정도 고졸) 태어날 모총출생아수는? 가정 : 타변수들의 값은 27세 부연령과 30세 모연령의 평균값을 입력 → 첫아이가 태어날 확률은 약 85% 정도임 → 2007년도 모총출생아수 출생방정식을 활용하여 계산한 결과 모총출생아수_i = 0.598log동거기간_i + 0.169다태아출산순위_i + 0.076모교육(고졸)_i + 0.104다태아여부(단태아)_i - 0.082log임신주수_i + 0.06출생아체중_i + 0.071log부연령_i + 0.049부모교육(고졸)_i 0.848 = 0.598log(3) + 0.169(1) + 0.076 + 0.104(1) - 0.082log(38.91) + 0.06(3.24) + 0.071log(27) + 0.049</p>