Journal of Korean
Data & Information Science Society
2003, Vol. 14, No.4 pp. 789~797

Test for Independence in Bivariate Weibull Model under Bivariate Random Censorship

Jang Sik Cho[1)]· Kil Ho Cho[2)] · Woo Dong Lee[3)]

Abstract

In this paper, we consider two components system which have bivariate weibull model with bivariate random censored data. We proposed large sample test for independence based on maximum likelihood estimator and relative frequency estimator, respectively. Also we derive asymptotic properties for the large sample tests and present a numerical study.

Keywords : Bivariate random censored data, Bivariate weibull model, Maximum likelihood estimator, Relative frequency estimator.

1. 서론

두 개의 부품으로 구성되어 있는 시스템에서 두 부품의 수명시간을 일반적으로 서로 종속적인 확률변수로 가정하는 것이 현실적인 경우가 많다. 이는 두 개의 부품을 가지는 시스템에서 한 부품이 고장나면 고장난 부품이 분담하던 부하가 정상부품에 전가되어 정상부품의 고장률에 영향을 미치기 때문이다. 예를 들어 사람의 양쪽 눈, 양쪽 콩팥 등 두 부품이 쌍으로 이루어진 시스템을 생각하면 각 쌍들의 수명은 서로 종속적인 관계에 있다.

이와 같이 부품의 수명이 서로 종속관계에 있는 시스템의 수명시간에 대한 모형으로서 Freund(1961), Marshall과 Olkin(1967), Block과 Basu(1974)등은 지수분포에 기초한 이변량 지수모형를 제안하면서 그 분포의 여러 가지 중요한 성질을 밝혔다.
한편, Lu(1989, 1992)와 Lu와 Bhattacharyya(1988, 1990)는 완전자료가 관찰되는 경우 Freund와 Marshall-Olkin의 이변량 지수모형을 확장한 모형으로서, 수명모형에서 폭

---

1) Corresponding Author. Associate Professor, Department of Statistical Information Science, Kyungsung University, Busan, 608-736, Korea.
E-mail : jscho@star.ks.ac.kr

2) Professor, Department of Statistics, Kyungpook National University, Daegu, 702-701, Korea.

3) Associate Professor, Faculty of Information and Science, Kyungsan University, Kyungpook, 712-240, Korea.

넓게 적용할 수 있는 이변량 와이블(BVW : bivariate Weibull) 모형을 소개하였다. Cho, Cha 그리고 Lee(2003), Cho, Kim 그리고 Kang(2003)은 완전자료 및 일변량 중단자료로 관측된 경우 시스템의 신뢰도를 추정하는 문제를 연구하였으며, Cho와 Cho(2003)은 제 1종 이변량 중단자료를 갖는 경우 신뢰도 추정문제를 연구하였다.

한편, 현실적으로 부품들의 수명이 실험자의 의도에 의해서 또는 실험환경의 제약에 의해서 두 부품의 수명이 임의로 중단되는 시간이 서로 다른 이변량 임의 중단된 자료(bivariate random censored data)로 관찰되는 경우가 많이 발생한다. 예를 들어 두 부품의 수명이 동일한 분포를 갖지 않는 경우 두 부품의 중단시간을 일변량 중단 모형보다 이변량 중단모형으로 하는 것이 타당하다. 여기서 두 부품의 중단시간이 미리 정해진 경우는 제 1종 이변량 중단모형이 되며, 두 부품의 임의중단 시간이 서로 같은 경우에는 일변량 임의중단 모형이 된다. 따라서 이변량 임의중단모형은 일변량 임의중단모형 및 제 1종 이변량 중단모형의 확장된 개념으로 가장 널리 쓰이는 중단 모형이다.

본 논문에서는 이변량 와이블모형을 따르는 두 개의 부품으로 구성된 시스템에서, 수명시간들이 이변량 임의 중단된 자료로 관찰되는 경우를 가정한다. 우도함수 및 상대도수에 기초한 모수의 추정량들을 각각 제안하고, 그 추정량의 근사적 성질을 밝힌다. 또한 최우추정량과 상대도수 추정량에 기초한 두 부품의 수명시간들의 독립성에 대한 대표본 검정법을 제안하고, 몬테칼로 모의실험을 통하여 제안된 두 검정법들을 각각 비교한다.

2. 개요

확률변수 (X, Y)를 두 부품의 수명시간이라 하자. 그리고 수명시간에 대한 분포를 모수 $(\zeta_1, \zeta_2, \zeta_3, \rho)$인 BVW 모형을 따른다고 가정한다면 결합확률밀도 함수는 식 (1)과 같이 주어진다.

$$
\begin{aligned}
&f(x, y; \zeta_1, \zeta_2, \zeta_3, \rho) \\
&= \begin{cases} \zeta_1(\zeta_2+\zeta_3)\rho^2 x^{\rho-1} y^{\rho-1} \exp[-\zeta_1 x^\rho - (\zeta_2+\zeta_3) y^\rho], & 0 < x < y < \infty, \\ \zeta_2(\zeta_1+\zeta_3)\rho^2 x^{\rho-1} y^{\rho-1} \exp[-(\zeta_1+\zeta_3) x^\rho - \zeta_2 y^\rho], & 0 < y < x < \infty, \\ \zeta_3 \rho x^{\rho-1} \exp[-\zeta x^\rho], & 0 < x = y < \infty, \end{cases}
\end{aligned} \tag{1}
$$

여기서 $\zeta_1, \zeta_2, \zeta_3, \rho > 0$ 이고, $\zeta = \zeta_1 + \zeta_2 + \zeta_3$이다.

그러면 (X, Y)에 대한 결합 신뢰도 함수는 식 (2)와 같이 주어지며,

$$
\begin{aligned}
\overline{F}(x, y) &= P(X > x, Y > y) \\
&= \exp[-(\zeta_1 x^\rho + \zeta_2 y^\rho + \zeta_3 \max(x, y)^\rho)],
\end{aligned} \tag{2}
$$

(X, Y) 에 대한 조건부 신뢰도 함수는 각각 식(3)과 식(4)로 주어진다.

$$\overline{F}_{X \mid Y=y}(x) = P(X > x \mid Y = y)$$

$$= \begin{cases} \exp(-\zeta_1 x^\rho), & y > x \\ \zeta_2(\zeta_2+\zeta_3)^{-1}\exp(-(\zeta_1+\zeta_3)x^\rho+\zeta_3 y^\rho), & x \geq y, \end{cases} \quad (3)$$

$$\overline{F}_{Y|X=x}(y) = P(Y > y \mid X = x)$$
$$= \begin{cases} \exp(-\zeta_2 y^\rho), & x > y \\ \zeta_1(\zeta_1+\zeta_3)^{-1}\exp(-(\zeta_2+\zeta_3)y^\rho+\zeta_3 x^\rho), & x \leq y. \end{cases} \quad (4)$$

위의 식 (1)에서 $\zeta_1 = \zeta_2$이면 두 부품의 수명시간에 대한 분포는 동일하며, $\zeta_3 = 0$이면 (X, Y)는 서로 독립인 와이블 분포를 따른다. 또한 $\rho = 1$인 경우, (X, Y)는 이변량 지수분포와 동일하다.

$j=1,2;\ k=1,2,3;\ i=1,2,\cdots,n$에 대하여 본 논문에서 사용되는 기호들을 소개하면 다음과 같다.

(1) (t_{x_i}, t_{y_i}) : i번째 시스템의 이변량 임의 중단시간.

(2) $G_{1i} = I(X_i > t_{x_i})$, $G_{2i} = I(Y_i > t_{y_i})$, $G^*_{ji} = 1 - G_{ji}$.

(3) $R_{1i} = I(X_i < Y_i)$, $R_{2i} = I(X_i > Y_i)$, $R_{3i} = I(X_i = Y_i)$, $R^*_{ki} = 1 - R_{ki}$.

(4) $\Theta = (\zeta_1, \zeta_2, \zeta_3, \theta_1, \theta_2)$.

그러면 이변량 임의중단된 자료가 관찰되는 경우, i번째 시스템에서 각각의 부품들에 대해서 다음과 같은 세 가지 경우가 발생할 수 있다.

(1) 두 개의 부품들이 모두 관찰되는 경우($G^*_{1i}G^*_{2i} = 1$).

(2) 하나의 부품만 관찰되고 다른 하나의 부품은 중단되는 경우($G_{1i}G^*_{2i} + G^*_{1i}G_{2i} = 1$).

(3) 두 개의 부품들이 모두 중단되는 경우($G_{1i}G_{2i} = 1$).

따라서 i번째 시스템의 수명시간 (x_i, y_i)은 다음과 같이 관찰된다.

$$(x_i, y_i) = \begin{cases} (x_i, y_i), & x_i < t_{x_i}\ ,\ y_i < t_{y_i} \\ (t_{x_i}, y_i), & x_i > t_{x_i}\ ,\ y_i < t_{y_i} \\ (x_i, t_{y_i}), & x_i < t_{x_i}\ ,\ y_i > t_{y_i} \\ (t_{x_i}, t_{y_i}), & x_i > t_{x_i}\ ,\ y_i > t_{y_i}\ , \end{cases} \quad (5)$$

여기서 임의중단시간 t_{x_i}와 t_{y_i} $(i=1,2,\cdots,n)$는 모수가 각각 (ρ, θ_1), (ρ, θ_2) 를 갖는 와이블 분포를 따르는 확률변수라 가정한다. 또한 생존분포는 각각 $\overline{G_1}(t_{x_i}:\rho, \theta_1)$과 $\overline{G_2}(t_{y_i}:\rho, \theta_2)$이고 확률밀도함수는 각각 $g_1(t_{x_i}:\rho, \theta_1)$과 $g_2(t_{y_i}:\rho, \theta_2)$이며, $f(x_i, y_i)$ 와 $g_1(t_{x_i}:\rho, \theta_1)$과 $g_2(t_{y_i}:\rho, \theta_2)$는 서로 독립이라 가정하자. 그러면 $t_{x_i} = t_{y_i}$인 경우는 일변량 임의 중단모형이 되고, 중단시간 t_{x_i}와 t_{y_i}이 미리 정해진 상수값이면 이변량 제1

종 중단모형이 된다.

따라서 우도함수는 식 (7)과 같이 계산될 수 있다.

$$
\begin{aligned}
L(\Theta) = \prod_{i=1}^{n} \{ & [f(x_i, y_i) \cdot \overline{G_1}(x_i;\rho,\theta_1) \cdot \overline{G_2}(y_i;\rho,\theta_2)]^{C_{1i}^{*}C_{2i}^{*}} \\
& \cdot [\overline{F}(x_i, y_i) \cdot g_1(x_i;\rho,\theta_1) \cdot g_2(y_i;\rho,\theta_{2i})]^{C_{1i}C_{2i}} \\
& \cdot [\overline{F}_{X\mid Y=y}(x_i) f_Y(y_i) \cdot g_1(x_i;\rho,\theta_1) \cdot \overline{G_2}(y_i;\rho,\theta_2)]^{C_{1i}C_{2i}^{*}} \\
& \cdot [\overline{F}_{Y\mid X=x}(y_i) f_X(x_i) \cdot \overline{G_1}(x_i;\rho,\theta_1) \cdot g_2(y_i;\rho,\theta_2)]^{C_{1i}^{*}C_{2i}} \}^{(R_{1i}+R_{2i}+R_{3i})} \\
= \ & \zeta_1^{D_1}\zeta_2^{D_2}\zeta_3^{D_3}(\zeta_1+\zeta_3)^{D_4}(\zeta_2+\zeta_3)^{D_5}\rho^{D_6} \cdot \theta_1^{D_6} \cdot \theta_2^{D_7} \\
& \cdot \prod_{i=1}^{n} \{ x_i^{(\rho-1)(C_{1i}+G_{1i}^{*})} y_i^{(\rho-1)(C_{2i}+G_{2i}^{*})(1-R_{3i}G_{1i}^{*})} \} \\
& \cdot \exp[-(\zeta_1+\theta_1)x_s - (\zeta_2+\theta_2)y_s - \zeta_3(x_s+y_s-t_s)],
\end{aligned} \tag{6}
$$

여기서 $f_X(x) = \rho(\zeta_1+\zeta_3)x_1^{\rho-1} \cdot \exp(-(\zeta_1+\zeta_3)x^{\rho})$,

$f_Y(y) = \rho(\zeta_2+\zeta_3)y^{\rho-1} \cdot \exp(-(\zeta_2+\zeta_3)y^{\rho})$,

$D_1 = \sum_{i=1}^{n}(R_{1i}G_{1i}^{*}G_{2i}^{*} + R_{2i}^{*}G_{1i}^{*}G_{2i})$, $D_2 = \sum_{i=1}^{n}(R_{2i}G_{1i}^{*}G_{2i}^{*} + R_{1i}^{*}G_{1i}G_{2i}^{*})$,

$D_3 = \sum_{i=1}^{n} R_{3i}G_{1i}^{*}G_{2i}^{*}$, $D_4 = \sum_{i=1}^{n} R_{2i}G_{1i}^{*}$, $D_5 = \sum_{i=1}^{n} R_{1i}G_{2i}^{*}$, $D_6 = \sum_{i=1}^{n} C_{1i}$, $D_7 = \sum_{i=1}^{n} C_{2i}$

$x_s = \sum_{i=1}^{n} x_i^{\rho}$, $y_s = \sum_{i=1}^{n} y_i^{\rho}$, $t_s = \sum_{i=1}^{n} \min(x_i, y_i)^{\rho}$.

그리고 $D_1, D_2, \cdots, D_5$는 확률변수이며, 이들의 기대값은 다음과 같이 계산되어 진다. 먼저 $E(D_1) = \sum_{i=1}^{n} E(R_{1i}G_{1i}^{*}G_{2i}^{*} + R_{2i}^{*}G_{1i}^{*}G_{2i})$이므로, 각각의 요소에 대해서 기대값을 계산하면 다음과 같다.

$$
\begin{aligned}
E(R_{1i}G_{1i}^{*}G_{2i}^{*}) &= P(0 < X_i < t_{x_i},\ 0 < Y_i < t_{y_i},\ X_i < Y_i) \\
&= \int_0^{t_{x_i}} \int_x^{t_{y_i}} \zeta_1(\zeta_2+\zeta_3)\rho^2 x^{\rho-1} y^{\rho-1} \exp(-\zeta_1 x^{\rho} - (\zeta_2+\zeta_3)y^{\rho})\,dydx \\
&= \zeta_1/\zeta - \zeta_1 \exp(-\zeta\, t^{\rho}{}_{x_i})/\zeta + \exp(-\zeta\, t^{\rho}{}_{y_i}) - \exp(-(\zeta_2+\zeta_3)\, t^{\rho}{}_{y_i}),
\end{aligned}
$$

$$
\begin{aligned}
E(R_{2i}^{*}G_{1i}^{*}G_{2i}) &= E(R_{1i}G_{1i}^{*}G_{2i} + R_{3i}G_{1i}^{*}G_{2i}) \\
&= P(X_i < t_{x_i},\ Y_i > t_{y_i},\ X_i < Y_i) + P(X_i < t_{x_i},\ Y_i > t_{y_i},\ X_i = Y_i) \\
&= \int_{t_{x_{i_{2i}}}}^{\infty} \int_0^{t_{x_i}} \rho^2 x_1^{\rho-1} x_2^{\rho-1} \zeta_1(\zeta_2+\zeta_3) \exp(-\zeta_1 x_1^{\rho} - (\zeta_2+\zeta_3)x_2^{\rho})\,dx_1 dx_2 \cdot I(t_{x_i} < t_{x_{i_{2i}}}) \\
&\quad + \int_{t_{y_i}}^{t_{x_i}} \zeta_3 \rho x_1^{\rho-1} \exp(-\zeta x_1^{\rho})\,dx_1 \cdot I(t_{y_i} < t_{x_i})
\end{aligned}
$$

$$= (1-\exp(-\zeta_1 t^{\rho}{}_{x_i}))\cdot\exp(-(\zeta_2+\zeta_3)\, t^{\rho}{}_{y_i})\cdot I(t_{x_i}<t_{y_i})$$
$$+\zeta_3(\exp(-\zeta\, t^{\rho}{}_{y_i})-\exp(-\zeta\, t^{\rho}{}_{x_i}))/\zeta\cdot I(t_{y_i}<t_{x_i}).$$

따라서,

$$E(D_1) = \sum_{i=1}^{n}\{\zeta_1/\zeta-\zeta_1\exp(-\zeta\, t^{\rho}{}_{x_i})/\zeta+\exp(-\zeta\, t^{\rho}{}_{y_i})-\exp(-(\zeta_2+\zeta_3)\, t^{\rho}{}_{y_i})$$
$$+(1-\exp(-\zeta_1 t^{\rho}{}_{x_i}))\cdot\exp(-(\zeta_2+\zeta_3)\, t^{\rho}{}_{y_i})\cdot I(t_{x_i}<t_{y_i})$$
$$+\zeta_3(\exp(-\zeta\, t^{\rho}{}_{y_i})-\exp(-\zeta\, t^{\rho}{}_{x_i}))/\zeta\cdot I(t_{y_i}<t_{x_i})\}.$$

같은 방법으로 계산을 한다면,

$$E(D_2)=\sum_{i=1}^{n}\{\zeta_2/\zeta-\zeta_2\exp(-\zeta\, t^{\rho}{}_{y_i})/\zeta+\exp(-(\zeta_1+\zeta_3)\, t^{\rho}{}_{x_i}-\zeta_2\, t^{\rho}{}_{y_i})$$
$$-\exp(-(\zeta_1+\zeta_3)\, t^{\rho}{}_{x_i})$$
$$+(1-\exp(-\zeta_2 t^{\rho}{}_{y_i}))\cdot\exp(-(\zeta_1+\zeta_3)\, t^{\rho}{}_{x_i})\cdot I(t_{y_i}<t_{x_i})$$
$$+\zeta_3(\exp(-\zeta\, t^{\rho}{}_{x_i})-\exp(-\zeta\, t^{\rho}{}_{y_i})))/\zeta\cdot I(t_{x_i}<t_{y_i})\},$$

$$E(D_3)=\sum_{i=1}^{n}\{(\zeta_3-\zeta_3\exp(-\zeta\min(t^{\rho}{}_{x_i},\ t^{\rho}{}_{y_i})))/\zeta\},$$

$$E(D_4)=\sum_{i=1}^{n}\{\zeta_2/\zeta-\zeta_2\exp(-\zeta\, t^{\rho}{}_{y_i})/\zeta+\exp(-(\zeta_1+\zeta_3)\, t^{\rho}{}_{x_i}-\zeta_2\, t^{\rho}{}_{y_i})$$
$$-\exp(-(\zeta_1+\zeta_3)\, t^{\rho}{}_{x_i})+[\exp(-\zeta_2\, t^{\rho}{}_{y_i})\cdot[1-\exp(-(\zeta_1+\zeta_3)\, t^{\rho}{}_{x_i})]$$
$$+(\zeta_1+\zeta_3)\cdot(\exp(-\zeta\, t^{\rho}{}_{x_i})-1)/\zeta]\cdot I(t_{x_i}>t_{y_i})\},$$

$$E(D_5)=\sum_{i=1}^{n}\{\zeta_1/\zeta-\zeta_1\exp(-\zeta\, t^{\rho}{}_{x_i})/\zeta+\exp(-\zeta\, t^{\rho}{}_{y_i})-\exp(-(\zeta_2+\zeta_3)\, t^{\rho}{}_{y_i})$$
$$+\zeta_1\exp(-\zeta\, t^{\rho}{}_{x_i})/\zeta-\exp(-(\zeta_2+\zeta_3)\, t^{\rho}{}_{y_i}-\zeta_1\, t^{\rho}{}_{x_i})\}.$$

본 논문에서는 고정된 ρ 값에 대해서만 초점을 맞추고자 한다. 그러면 최우추정량을 구하기 위한 로그-우도함수는 식 (7)과 같이 주어진다.

$$\log L(\Theta)=D_1\log\zeta_1+D_2\log\zeta_2+D_3\log\zeta_3+D_4\log(\zeta_1+\zeta_3)+D_5\log(\zeta_2+\zeta_3)$$
$$+D_6\log(\theta_1)+D_7\log(\theta_2)$$
$$+\sum_{i=1}^{n}\{(\rho-1)(C_{1i}+G^{*}_{1i})\log(x_i)+(\rho-1)(C_{2i}+G_{2i})(1-R_{3i}G^{*}_{1i})\log(y_i)\}$$
$$-(\zeta_1+\theta_1)x_s-(\zeta_2+\theta_2)y_s-\zeta_3(x_s+y_s-t_s). \tag{7}$$

따라서 우도방정식은 각각 식 (8) - 식 (12)과 같다.

$$\frac{\partial}{\partial\zeta_1}\log L(\Theta)=\frac{D_1}{\zeta_1}+\frac{D_4}{\zeta_1+\zeta_3}-x_s=0. \tag{8}$$

$$\frac{\partial}{\partial\zeta_2}\log L(\Theta)=\frac{D_2}{\zeta_2}+\frac{D_5}{\zeta_2+\zeta_3}-y_s=0. \tag{9}$$

$$\frac{\partial}{\partial \zeta_3} \log L(\Theta) = \frac{D_3}{\zeta_3} + \frac{D_4}{\zeta_1 + \zeta_3} + \frac{D_5}{\zeta_2 + \zeta_3} - (x_s + y_s - t_s) = 0. \tag{10}$$

$$\frac{\partial}{\partial \theta_1} \log L(\Theta) = \frac{D_6}{\theta_1} - x_s = 0. \tag{11}$$

$$\frac{\partial}{\partial \theta_2} \log L(\Theta) = \frac{D_7}{\theta_2} - y_s = 0, \tag{12}$$

모수 Θ에 대한 최우추정량 $\widehat{\Theta}$은 위의 우도방정식 (8)-(12)을 뉴턴-랍슨과 같은 수치 해석적 방법에 의해 얻을 수 있다. 그리고 (θ_1, θ_2)에 대한 최우추정량은 식 (11)와 식 (12)에 의해서 $\widehat{\theta}_1 = D_6/x_s$, $\widehat{\theta}_2 = D_7/y_s$로 얻을 수 있다.

따라서 $\sqrt{n}\,(\,\widehat{\underset{\sim}{\zeta}}^T - \underset{\sim}{\zeta}^T)$의 분포는 표본의 크기가 증가하면서 근사적으로 평균벡터가 영이고 분산-공분산행렬이 $I^{-1}(\underset{\sim}{\zeta})$인 다변량 정규분포를 따름을 알 수 있다. 여기서 $\underset{\sim}{\zeta} = (\zeta_1, \zeta_2, \zeta_3)$, $\widehat{\underset{\sim}{\zeta}} = (\widehat{\zeta}_1, \widehat{\zeta}_2, \widehat{\zeta}_3)$.

그리고 피셔의 정보행렬은 다음과 같이 계산할 수 있다.

$I(\underset{\sim}{\zeta}) = ((I_{ij}))$, 여기서 $I_{ij} = E\left[-\frac{\partial^2}{\partial \zeta_i \partial \zeta_j} \log L(\underset{\sim}{\zeta})\right]$; $i, j = 1, 2, 3$이며,

$$I_{11} = E(D_1)/\zeta_1^2 + E(D_4)/(\zeta_1 + \zeta_3)^2,\ \ I_{12} = 0,\ \ I_{13} = E(D_4)/(\zeta_1 + \zeta_3)^2,$$
$$I_{22} = E(D_2)/\zeta_2^2 + E(D_5)/(\zeta_2 + \zeta_3)^2, I_{23} = E(D_5)/(\zeta_2 + \zeta_3)^2,$$
$$I_{33} = E(D_3)/\zeta_3^2 + E(D_4)/(\zeta_1 + \zeta_3)^2 + E(D_5)/(\zeta_2 + \zeta_3)^2.$$

3. 독립성에 대한 대표본 검정

이 장에서는 최우추정량 $\widehat{\zeta}_3$과 상대도수에 기초한 추정량 $m_3 = \sum_{i=1}^{n} R_{3i}$ 를 이용하여 두 부품의 수명시간들에 대한 독립성에 대한 대표본 검정을 하고자 한다.
2장에서 알 수 있는 바와 같이 독립성 검정을 위한 귀무가설은 $H_o: \zeta_3 = 0$으로 설정할 수 있다. 따라서 최우추정량 $\widehat{\zeta}_3$와 상대도수 m_3에 기초하여 $H_o: \zeta_3 = 0$에 대한 검정통계량을 구성하기로 한다.

먼저 최우추정량 $\widehat{\zeta}_3$에 기초한 독립성에 대한 대표본 검정을 생각해 보자. $\widehat{\zeta}_3$에 대한 정확한 분포는 알 수 없지만, 2장의 결과를 이용하면 $\widehat{\zeta}_3$의 분포는 근사적으로 평균이 ζ_3이고, 분산이 I^{33}/n인 정규분포를 따름을 알 수 있다. 그러나 I^{33}의 값이 미지의 모수 $(\zeta_1, \zeta_2, \zeta_3)$의 값에 의존하기 때문에 아래와 같이 근사적 정규분포를 따르는 검정통계량을 구성할 수 있다.

$$\phi_{MLE}(X, Y) = \sqrt{n} \cdot \widehat{\zeta}_3 / \sqrt{\widehat{I}^{33}}, \tag{13}$$

여기서 검정통계량 $\phi_{MLE}(X, Y)$는 근사적으로 표준정규분포를 따르며, $\widehat{I}^{ij}$는 I^{ij}속에 내재되어 있는 미지의 모수 값 대신 최우추정량을 대입한 값이다.

따라서 가설 $H_1: \zeta_3 > 0$에 대하여 유의수준 α에서 H_o을 기각하기 위해서는 다음의 조건이 만족되어야 한다.

$$\phi_{MLE}(X, Y) = \sqrt{n} \cdot \widehat{\zeta}_3 / \sqrt{\widehat{I}^{33}} \ > \ z_{1-\alpha}, \tag{14}$$

여기서 z_α는 표준정규분포의 오른쪽 꼬리 면적이 α가 되는 값을 의미한다.

다음으로 상대도수 m_3에 기초해서 독립성에 대한 대표본 검정을 생각해 보자. m_3의 분포는 모수가 ($n, \zeta_3/\zeta$)인 이항분포를 따르며, 대표본에 기초한 검정통계량은 다음과 같이 구성할 수 있다.

$$\phi_{RF}(X, Y) = \sqrt{n} m_3 / \sqrt{m_3(n - m_3)}, \tag{15}$$

여기서 검정통계량 $\phi_{RF}(X, Y)$는 근사적으로 표준정규분포를 따른다. 따라서 가설 $H_1: \zeta_3 > 0$에 대해서 유의수준 α에서 귀무가설을 기각하기 위해서는 다음의 조건이 만족되어야 한다.

$$\phi_{RF}(X, Y) = \sqrt{n} m_3 / \sqrt{m_3(n - m_3)} \ > \ z_{1-\alpha}. \tag{16}$$

4. 모의실험

이 장에서는 난수생성기에 의해 생성된 가상자료를 이용하여 3장에서 구성한 두 가지의 대표본에 입각한 독립성 검정법을 비교하고자 한다. 먼저 본 논문에서는 수명분포에 대한 이변량 와이블 분포에서는 모수 값으로 $\rho = 2.0$ 및 ($\zeta_1 = 1.2$, $\zeta_2 = 1.4$, $\zeta_3 = 0.2$)을 설정하였으며, 이변량 와이블 분포로부터 표본크기 30인 두 부품의 수명시간에 해당하는 난수를 발생시켰다. 설정된 모수의 값으로부터 대립가설 $H_1: \zeta_3 > 0$이 타당함을 알 수 있다. 이변량 임의 중단시간에 대한 분포로서 모수가 각각 $(\rho, \theta_1) = (2,\ 0.2)$, $(\rho, \theta_2) = (2,\ 0.4)$인 와이블 분포를 이용하였다. 관측된 자료는 아래 [표 1]과 같으며, 여기서 *는 임의중단된 자료를 나타낸다.

[표 1] 이변량 와이블 분포로부터 생성된 자료

i	x_i	y_i	i	x_i	y_i	i	x_i	y_i
1	1.0335	0.8088	11	0.6406	0.2766	21	0.7731	1.0065
2	0.9801*	0.3119	12	0.3076	0.5736	22	0.1589	0.6147
3	0.5192	0.0867	13	0.1689	0.7503	23	0.8451	0.5704
4	0.3850	0.4267	14	0.6444*	0.6704	24	0.2264	0.3176
5	1.1681	0.3354	15	0.5418	0.8880	25	0.9226	0.2738
6	0.9166	0.4641*	16	0.7040	0.4692	26	1.2092*	0.7345
7	1.1004	0.7124	17	0.3113	0.3113	27	0.5397	0.5128
8	1.3180	1.1582	18	0.6545	0.8104	28	1.2486	1.2103
9	0.6628	1.4367*	19	0.9478*	0.9725	29	0.4028	0.4363
10	0.5171	0.6498*	20	0.3559	0.4606	30	0.8222	0.9761*

[표 1]의 자료로부터 모수에 대한 최우추정량은 $\widehat{\zeta}_1=1.3281$, $\widehat{\zeta}_2=1.5774$, $\widehat{\zeta}_3=0.1355$와 같이 계산되며, m_3에 대한 값은 1로 계산된다.

따라서 식 (12)와 식 (14)에 대한 검정통계량 값과 유의확률을 계산하면 아래 [표 2]와 같다.

[표 2] 두 가지 검정법의 검정통계량과 유의확률

검정통계량 / 계산결과	$\phi_{MLE}(X, Y)$	$\phi_{RF}(X, Y)$
검정통계량	6.5283	1.0171
유의확률	0.0000	0.1545

[표 2]로부터 최우추정량 $\widehat{\zeta}_3$에 기초한 검정통계량 $\phi_{MLE}(X, Y)$의 값은 6.5283으로서, 유의수준 0.001에서도 귀무가설 $H_o: \zeta_3=0$을 기각할 수 있다. 그러나 상대도수 m_3에 기초한 검정통계량 $\phi_{RF}(X, Y)$의 값은 1.0171로서, 유의수준 0.10에서 조차도 귀무가설 $H_o: \zeta_3=0$을 기각할 수 없음을 알 수 있다. 따라서 위의 데이터에서는 독립성에 대한 검정법은 최우추정량 $\widehat{\zeta}_3$에 기초한 검정통계량이 m_3에 기초한 검정법에 비해서 다소 우수함을 알 수 있다.

모수와 임의중단분포 및 표본크기에 대한 다양한 조합에 대해서도 본 논문에서 제안한 두 가지 검정법의 비교와 다른 이변량 수명분포로 확장해서 다양한 검정법에 대한 연구는 향후 과제로 남겨두고자 한다.

참고문헌

1. Block, H. W. and Basu, A. P. (1974). A Continuous Bivariate Exponential Extension, Journal of the American Statistical Association, 69, 1031-1037.

2. Cho, J. S., Cha, Y. J. and Lee, J. M. (2003). Estimation of System Reliability in Bivariate Weibull Distribution, Journal of The Korean Data Analysis Society, Vol. 5(1), 11-16.
3. Cho, J. S. and Cho, K. H. (2003). Reliability for Series System in Bivariate Weibull Model under Bivariate Type I Censorship, Journal of The Korean Data & Information Science Society, Vol. 14(3), 571-578.
4. Cho, J. S., Kim, H. J. and Kang, S. G. (2003). Estimation of Reliabilities in Bivariate Weibull Model Under Random Censorship, Journal of The Korean Data Analysis Society, Vol. 5(2), 145-152.
5. Freund, J. E. (1961). A bivariate extension of the exponential distribution, Journal of the American Statistical Association, 971-977.
6. Lu, J.C. (1989). Weibull extensions of the Freund and Marshall-Olkin bivariate exponential models, IEEE Transactions on Reliability, 38, 615-619.
7. Lu, J.C. (1992). Bayes parameter estimation for the bivariate Weibull model of Marshall-Olkin for censored data, IEEE Transactions on Reliability, 41(4), 608-615.
8. Lu, J.C. and Bhattacharyya, G.K. (1988). Some bivariate extensions of the Weibull distribution, Technical Report No. 821, Department of Statistics, University of Wisconsin, Madison.
9. Lu, J.C. and Bhattacharyya, G.K. (1990). Some new constructions of bivariate Weibull models, Annals Institute Statistical Mathematics, 42, 543-559.
10. Marshall, A.W. and Olkin, I. (1967). A Multivariate Exponential Distribution, Journal of the American Statistical Association, 62, 30-44.

[2003년 9월 접수, 2003년 11월 채택]