

이산형 반응변수에서 오류 분배율 함수를 적용한 집단축차 검정

김동욱^{1,a}, 남진현^a

^a성균관대학교 통계학과

요약

본 논문에서는 중간분석에서 사용되는 집단축차 검정법으로 이산형 반응변수인 경우, 오류 분배율 함수를 적용한 집단축차 검정법을 제안한다. 특히 제 1종 오류와 제 2종 오류를 모두 적용한 집단축차 검정법을 제안하며, 기존의 오류 분배율 함수를 포함하는 새로운 오류 분배율 함수를 제안한다. 반응변수가 이산형인 경우 정확한 크기 α 검정을 할 수 없으므로 각 검정단계에 사용될 오류율을 분배하는 대신 각 검정단계까지 사용되어야 할 누적 오류율을 이용한다. 오류 분배율 함수를 적용한 집단축차 검정은 기존의 집단축차 검정보다 빠른 연산과 유연한 검정이 가능하다는 장점을 지니고 있으며, 본 논문에서 제시된 오류 분배율 함수를 이용해 특성을 비교한다.

주요어: 이산형변수, 오류 분배율 함수, 집단축차, 중간분석.

1. 서론

임상시험을 시행함에 있어 계획된 시간에 단 한 번의 분석으로 치료법의 효과 등을 검정하는 것은 통계적 측면에서 볼 때 비교적 간단하고 용이하다. 그러나 연구의 합리적인 운용과 환자의 안전을 위해서 정기적인 중간 점검(interim monitoring)이 필요하다. 중간 점검을 통해 임상 시험의 진행 여부를 결정하며, 그 기준으로 중간분석(interim analysis)이 이용된다. 중간분석에서 얻어진 처리군들의 차이에 대한 평가는 매우 중요하므로 중간분석 시 합리적인 통계적 기준이 필요하다. 따라서 연구 설계 단계에서부터 이러한 중간 점검을 고려한 통계적 모형을 사용해야 하며, 임상시험의 결과로부터 생성된 자료를 효율적으로 통제하는 중간 점검 방법으로 집단축차 검정법(group sequential test)을 사용한다.

집단축차 검정법은 일정한 시점마다 그때까지 누적되어 온 자료를 통해 계획된 분석을 반복 시행하는 검정 방법이다. 이 방법은 특정 임상시험에 관하여 사전에 분석 시점, 분석 횟수, 최종 분석시기 등의 다양한 결정을 해야 한다. 집단축차 검정에서 가장 큰 관심사는 주어진 조건하에서 검정을 시행할 때마다 증가하는 제 1종 오류율을 요구하는 제 1종 오류의 크기로 어떻게 통제(control)하는가이다. 이에 관해 Pocock (1977)과 O'Brien과 Fleming (1979)은 반응변수가 연속형인 경우 각 분석시점 간에 추가된 관측치의 수가 동일하다는 가정 하에 제 1종 오류를 만족하는 기각역을 제시하였다. Fleming (1982)은 반응변수가 이산형인 경우 정규근사를 이용한 기각역을 제시했으며, Chang 등 (1987)은 최강력 검정을 이용해 평균 표본수(average sample number; ASN)를 최소화하는 기각역을 제안했다. 이러한 집단축차 검정법은 제 1종 오류를 잘 통제하지만, 분석 시점, 분석 횟수, 최종 분석시기 등 많은 제약이 따른다. 이러한 제약을 완화하기 위해 오류 분배율 함수(error spending rate function)가 논의되었다. Lan과 DeMets (1983)는 반응변수가 연속형인 경우 제 1종 오류 분배율 함수(type I error spending rate function)를 제안함으로써 많은 제약을 완화시켰으며, 제 1종 오류 분배율 함수의 특성에 의해 각 검정 시점마다 누적된 제 1종 오류의 크기가 주어진 제 1종 오류의 크기를 넘지 않는 범위 내에서 제 1종 오

¹ 교신저자: (110-745) 서울시 중로구 명륜동 3가 53, 성균관대학교 통계학과, 교수. E-mail: dkim@skku.edu

류를 원하는 만큼 분배할 수 있게 되었다. 이에 Chang 등 (1998)은 Lan과 DeMets (1983)의 방법을 확장하여 제 1종 오류뿐만 아니라 제 2종 오류도 분배하는 방법을 제안하였다.

본 논문은 반응변수가 이산형일 경우 제 1종 오류 분배율 함수와 제 2종 오류 분배율 함수를 모두 적용한 집단축차 검정법을 제안한다. 이는 제 1종 오류 분배율 함수를 적용한 집단축차 검정법으로도 쉽게 적용이 가능하다. 아울러 기존에 제시된 오류 분배율 함수를 포괄하는 새로운 오류 분배율 함수를 제안하며 이를 통해 반응변수가 이산형인 경우, 특히 이항분포인 경우 오류 분배율 함수에 따른 특성을 연구한다.

제 2절에서는 집단축차 검정법을 소개하며 제 3절에서는 반응변수가 이산형인 경우 제 1종, 제 2종 오류 분배율 함수를 적용한 집단축차 검정법을 소개한다. 제 4절에서는 새로운 오류 분배율 함수를 제안하며, 오류 분배율 함수의 특성을 살펴본다. 제 5절에서 결론을 맺는다.

2. 집단축차 검정법

임상시험이 진행되는 도중에 처리군들 간의 데이터를 점검하는 가장 중요한 이유로 효과가 작거나 부작용이 큰 처리군에 더 이상 환자를 할당하지 않도록 하는 윤리적인 면과 효과가 없는 처리에 대해 더 이상 임상시험을 지속하지 않는 경제적인 면을 들 수 있다. 따라서 중간분석에서 얻어진 처리군들에 대한 평가는 매우 중요한 일이며, 중간분석 결과를 어떻게 다루고 해석하는가에 대한 객관적이고 합리적인 기준을 마련하는 것이 필요하다.

중간분석을 실시할 때마다 유의성 검정을 하는데 있어서 가장 큰 문제는, 실제로 두 처리군의 효과가 같더라도 일정한 유의수준에서 차이가 있다고 결론을 내릴 가능성이 한번만 검정하는 경우보다 훨씬 크다는 것이다 (Armitage 등, 1969). 원하는 수준의 제 1종 오류를 얻기 위해서 축차 검정법이 개발되었으나, 이 방법은 새로운 데이터가 얻어질 때마다 분석해야 하는 제약이 있어서 임상시험에 적용하기에는 많은 어려움이 따른다. 때문에 보다 현실적인 방법으로 임상시험 중 몇 차례 정해진 수의 계획된 분석을 실시하는 집단축차 검정법이 연구되었다 (Pocock, 1977; O'Brien과 Fleming, 1979; Schultz 등, 1973; Fleming, 1982; Chang 등, 1987).

이와 같은 방법들은 매우 유용하지만, 제약이 많다. 기본적으로 검정의 총 시행 횟수나 시행 시점, 각 검정 시 사용되는 그룹의 크기 등이 사전에 정의되어야 하는데, 실제적으로 이러한 조건은 몇몇 문제를 발생시킨다. 임상시험을 하는 동안에 임상시험의 의사결정 그룹에 의해 데이터를 검토하는 횟수를 늘리거나 혹은 줄여야 하는 상황이 발생할 수도 있으며, 지원자의 참여가 처음에 계획했던 것보다 느려지는 경우에 불가피하게 분석 횟수를 증가시켜야 하는 문제가 생길 수도 있다. Lan과 DeMets (1983)는 이러한 문제의 해결책으로 반응변수가 연속형인 경우에 $\alpha^*(0) = 0$ 이고 $\alpha^*(1) = \alpha$ 를 만족하는 비감소 함수인 제 1종 오류 분배율 함수를 적용한 집단축차 검정법을 제안하였다. 이에 Kim과 DeMets (1987)는 비대칭 양측 검정으로 확장을 하였고, Chang 등 (1998)은 제 1종 오류 분배율 함수와 제 2종 오류 분배율 함수 모두 이용함으로써 제 1종 오류뿐만 아니라 검정력까지 고려할 수 있게 하였다.

귀무가설 하에서 $S_k, a_k, r_k, k = 1, \dots, K$ 를 k 번째 중간분석에서 각각 표준화된 검정통계량, 채택점, 기각점이라고 하자. k 번째 중간분석에서 귀무가설을 기각할 확률을 $P_k, k = 1, \dots, K$ 라고 하면 P_k 는 식 (2.1)과 같이 정의할 수 있다.

$$P\{|S_1| \leq r_1, \dots, |S_{k-1}| \leq r_{k-1}, |S_k| > r_k | H_0\} = P_k, \quad \sum_{k=1}^K P_k = \alpha. \quad (2.1)$$

식 (2.1)에서 $P_k = \alpha^*(t_k) - \alpha^*(t_{k-1})$ 로 두면, 즉

$$P\{|S_i| \leq r_i, i = 1, \dots, k-1, |S_k| > r_k | H_0\} = \alpha^*(t_k) - \alpha^*(t_{k-1}) \quad (2.2)$$

이런 제 1종 오류 분배율 함수를 적용한 집단축차 검정법이 된다. 여기서 정보시점(information time) t_k 는 k 번째 중간분석에서 정보시간의 비율을 나타낸다. 이를 확장하여 제 1종, 제 2종 오류 분배율 함수를 적용한 집단축차 검정법은 다음과 같다.

$$P(a_1 \leq |S_1^*| \leq r_1, \dots, a_{k-1} \leq |S_{k-1}^*| \leq r_{k-1}, |S_k^*| > r_k | H_0) = \alpha^*(t_k) - \alpha^*(t_{k-1}). \quad (2.3)$$

$$P(a_1 \leq |S_1^*| \leq r_1, \dots, a_{k-1} \leq |S_{k-1}^*| \leq r_{k-1}, |S_k^*| < a_k | H_1) = \beta^*(t_k) - \beta^*(t_{k-1}). \quad (2.4)$$

오류 분배율 함수를 적용한 집단축차검정은 적절한 오류 분배율 함수를 선택함으로써 임상시험의 성격에 맞는 기각역을 설정할 수 있으며 검정 시점에 제약을 받지 않기 때문에 유연한(flexible) 검정이 가능하다.

3. 이산형 반응변수에 오류 분배율 함수 적용

3.1. 제 1종과 제 2종 오류 분배율 함수 적용

각 검정 단계에서 추가되는 환자 수를 n_1, \dots, n_K 라 하고 $N_k = \sum_{i=1}^k n_i$, $N = \sum_{i=1}^K n_i$, $k = 1, \dots, K$ 라 하자. s_i 를 n_i 환자 중 성공한 개체수로 나타내면 k 번째 단계에서 성공한 개체수의 총 누적합 S_k 는 $S_k = \sum_{i=1}^k s_i$, $k = 1, \dots, K$ 가 된다. 이때 S_k 는 모수가 (N_k, p) 인 이항분포를 따르게 된다.

검정하려는 가설이 $H_0 : p \leq p_0$ vs. $H_1 : p \geq p_1$, ($p_0 < p_1$)일 때, 채택점을 (a_1, a_2, \dots, a_K) , 기각점을 (r_1, r_2, \dots, r_K) 라 하면, 채택점과 기각점의 관계는 다음과 같다.

$$r_k - 1 > a_k, \quad k = 1, \dots, K-1 \quad \text{그리고} \quad a_K = r_K - 1.$$

검정단계 k 에서 추가되는 환자수는 n_k 이고, 성공한 개체수를 s_k 라고 할 때 $s_k = v$ 의 확률은 특정한 p 의 값이 주어질 때 식 (3.1)과 같이 이항 분포로 나타내어진다.

$$B_k(v, p) = \binom{n_k}{v} p^v (1-p)^{n_k-v}, \quad k = 1, \dots, K. \quad (3.1)$$

또한, $k-1$ 번째 검정까지 보류를 하고 k 번째 검정에서 $S_k = m$ 일 확률은 식 (3.2)와 같다.

$$C_k(m, p) = \sum_{j=E}^F C_{k-1}(j, p) B_k(m-j, p), \quad (3.2)$$

여기서 $E = \max(a_{k-1} + 1, m - n_k)$ 이고 $F = \min(r_{k-1} - 1, m)$ 이다. 특히, 첫 번째 단계에서 $S_1 = m$ 일 확률은 $C_1(m, p) = B_1(m, p)$ 이다.

k 번째 검정 단계에서 H_0 를 채택할 확률은

$$L_k(p) = \sum_{m=a_{k-1}+1}^{a_k} C_k(m, p)$$

이고, 마지막 검정 단계까지에서 H_0 를 채택할 누적 확률은 식 (3.3)과 같다.

$$L(p) = \sum_{k=1}^K L_k(p). \quad (3.3)$$

마찬가지로 k 번째 검정 단계에서 H_0 를 기각할 확률은

$$R_k(p) = \sum_{m=r_k}^{r_{k-1}-1+n_k} C_k(m, p)$$

이고, 마지막 검정 단계까지에서 H_0 를 기각할 누적 확률은 식 (3.4)와 같다.

$$R(p) = \sum_{k=1}^K R_k(p). \quad (3.4)$$

이때, ASN은 다음과 같다 (Schultz 등, 1973).

$$ASN = \sum_{k=1}^K N_k [L_k(p) + R_k(p)].$$

오류 분배율 함수를 이용하기 위해 정보 시점 t_k , $k = 1, \dots, K$ 는 $t_k = N_k/N$, $0 \leq t_k \leq 1$ 로 정의한다 (Lan과 DeMets, 1984).

집단축차 검정방법은 다음과 같다.

만약 $S_k \leq a_k$ 이면, 피험자 모집을 중단하고 $H_1 : p \geq p_1$ 를 기각한다.

만약 $S_k \geq r_k$ 이면, 피험자 모집을 중단하고 $H_0 : p \leq p_0$ 를 기각한다.

만약 $a_k < S_k < r_k$ 이면, $k+1$ 단계에서 검정을 계속한다.

최대 표본 크기 N , 요구되는 전체 유의 수준 α , 검정력 $1 - \beta$ 그리고 제 1종, 제 2종 오류 분배율 함수 $\alpha^*(t_k)$, $\beta^*(t_k)$ 가 주어졌을 때, 기각점과 채택점을 구하기 위해 t_k ($0 < t_1 < \dots < t_K = 1$) 시점에서 검정 통계량 S_k 를 구한다. 검정 통계량 S_k 는 이산형이기 때문에 각 검정 단계에서 사용되는 제 1종, 제 2종 오류율을 만족하는 정확한 임계값을 찾을 수 없으며, 이로 인해 각 검정 시점 t_k 에서 사용되어야 할 오류율들을 전부 사용하지 못하는 문제점이 발생한다. 따라서 이러한 문제를 해결하기 위해 각 검정단계에서 사용되는 오류율이 아닌 각 검정단계까지 사용되어야 할 누적 오류율을 통해 임계값을 정하도록 한다.

첫 번째 기각점은 식 (3.5)를 만족하는 r_1 중 가장 작은 r_1 을 선택하며, 첫 번째 채택점은 식 (3.6)을 만족하는 a_1 중 가장 큰 a_1 을 선택한다.

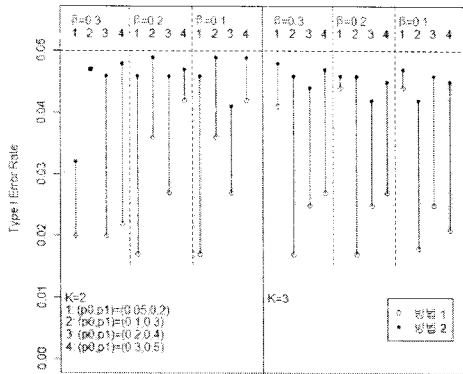
$$P\{S_1 \geq r_1 | H_0\} \leq \alpha^*(t_1). \quad (3.5)$$

$$P\{S_1 \leq a_1 | H_1\} \leq \beta^*(t_1). \quad (3.6)$$

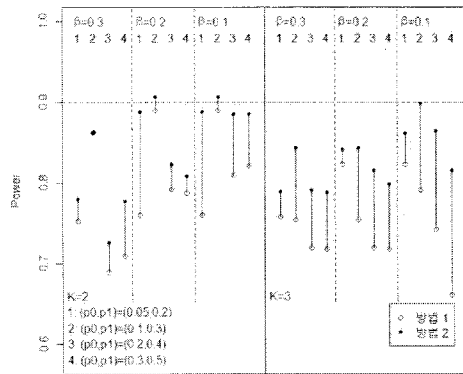
첫 번째 채택점과 첫 번째 기각점 (a_1, r_1)을 통해 실제 사용된 제 1종 오류율과 제 2종 오류율은 다음과 같다.

$$R_1(p_0) = \sum_{m=r_1}^{n_1} C_1(m, p_0).$$

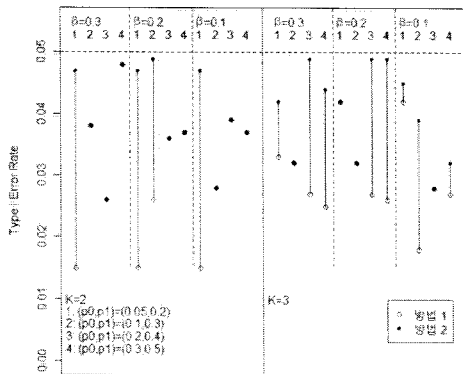
$$L_1(p_1) = \sum_{m=0}^{a_1} C_1(m, p_1).$$



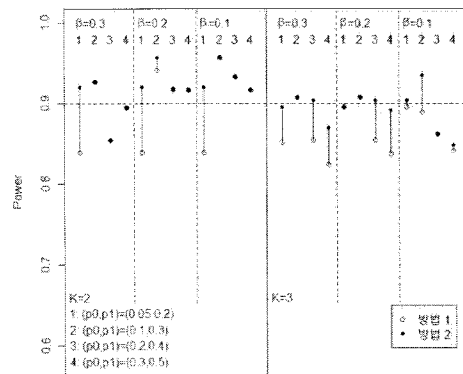
(a) $(p_0, p_1), \beta, K$ 의 변화에 따른 제 1종 오류율



(b) $(p_0, p_1), \beta, K$ 의 변화에 따른 검정력



(c) $(p_0, p_1), \beta, K$ 의 변화에 따른 제 1종 오류율



(d) $(p_0, p_1), \beta, K$ 의 변화에 따른 검정력

그림 1: $(p_0, p_1), \beta, K$ 의 변화에 따른 제 1종 오류율과 검정력 비교

정보시점 $t_k, k = 2, \dots, K - 1$ 일 때 k 번째 기각점은 식 (3.7)을 만족하는 r_k 중 가장 작은 r_k 를 선택하며, k 번째 채택점은 식 (3.8)을 만족하는 a_k 중 가장 큰 a_k 를 선택한다.

$$P\{a_j < S_j < r_j, j = 1, \dots, k - 1; S_k \geq r_k | H_0\} \leq \alpha^*(t_k) - \sum_{i=1}^{k-1} R_i(p_0). \quad (3.7)$$

$$P\{a_j < S_j < r_j, j = 1, \dots, k - 1; S_k \leq a_k | H_1\} \leq \beta^*(t_k) - \sum_{i=1}^{k-1} L_i(p_1). \quad (3.8)$$

K 번째 단계에서 기각점은 식 (3.7)을 만족하는 r_K 중 가장 작은 r_K 을 선택하며, 채택점 a_K 는 $a_K = r_K - 1$ 이 된다.

그림 1은 $\alpha = .05$ 인 경우 $(p_0, p_1) = \{(0.05, 0.2), (0.1, 0.3), (0.2, 0.4), (0.3, 0.5)\}, \beta = \{0.3, 0.2, 0.1\}, K = 2, 3$ 을 변화시키며 각 검정단계에서 사용되는 오류율을 이용하여 임계값을 설정하는 <방법 1>과 각 검정단계까지 사용되어야 할 누적 오류율을 이용하여 임계값을 설정하는 <방법 2>의 제 1종 오류율과 검정력을 비교한 것이다. 그림 1의 (a)와 (b)는 식 (4.5)의 $a = 1, b = 3$ 인 결과이며, (c)와 (d)는 식

(4.5)의 $a = 3, b = 1$ 인 결과이다. 그림 1에서 알 수 있듯이 <방법 2>가 <방법 1>보다 주어진 제 1종 오류율을 더 활용하며 검정력 또한 <방법 2>가 훨씬 높다는 것을 알 수 있다. ASN을 비교해도 <방법 1>보다 <방법 2>가 더 작게 나온다. 따라서 <방법 1>보다 <방법 2>를 사용하는 것이 타당하다.

3.2. 제 1종 오류 분배율 함수만의 적용

3.1절에서 조기 기각역과 조기 채택역을 모두 고려한 검정법을 제시하였다. 다음으로 조기 기각역만 고려한 즉, 제 1종 오류 분배율 함수만 적용한 방법을 제시한다. k 번째 단계 $k = 1, \dots, K-1$, 검정에서 조기 채택역은 존재하지 않고, 마지막 검정 단계에서만 채택점이 $a_K = r_K - 1$ 인 검정 방법은 다음과 같다.

만약 $S_k \geq r_k$ 이면, 표본 모집을 중단하고 H_0 을 기각한다.

만약 $S_k < r_k$ 이면, $k+1$ 번째 검정을 계속한다.

이 방법의 경우, $S_k = m$ 의 확률은 식 (3.9)와 같다.

$$C_k(m, p) = \sum_{j=0}^{r_{k-1}-1} C_{k-1}(j, p) B_k(m-j, p). \quad (3.9)$$

기각점을 구하는 방법은 다음과 같다. 첫 번째 기각점은 식 (3.10)을 만족하는 r_1 중 가장 작은 r_1 을 선택한다.

$$P\{S_1 \geq r_1 | H_0\} \leq \alpha^*(t_1). \quad (3.10)$$

첫 번째 기각점 r_1 을 통해 실제 사용된 제 1종 오류율은 다음과 같다.

$$R_1(p_0) = \sum_{m=r_1}^{n_1} C_1(m, p_0).$$

$k = 2, \dots, K-1$ 일 때 k 번째 기각점은 식 (3.11)을 만족하는 r_k 중 가장 작은 r_k 를 선택한다.

$$P\{a_j < S_j < r_j, j = 1, \dots, k-1; S_k \geq r_k | H_0\} \leq \alpha^*(t_k) - \sum_{i=1}^{k-1} R_i(p_0). \quad (3.11)$$

K 번째 단계에서 기각점은 식 (3.11)을 만족하는 r_K 중 가장 작은 r_K 를 선택하며, 채택점 a_K 는 $a_K = r_K - 1$ 이 된다.

4. 오류 분배율 함수의 비교

기존의 오류 분배율 함수는 ρ 족 (Kim과 DeMets, 1987)과 γ 족 (Hwang 등, 1990)의 형태로 표현 가능하다 (Barber와 Jennison, 2002; Anderson, 2006). ρ 족은 다음과 같은 형태이다.

$$\alpha^*(t, \rho) = \alpha \min(1, t^\rho) \quad (4.1)$$

$$\beta^*(t, \rho) = \beta \min(1, t^\rho), \quad \rho > 0.$$

ρ 족은 ρ 값에 따라 특성이 다른 여러 가지의 오류 분배율 함수를 만들 수 있고, Kim과 DeMets (1987)는 $\rho = 3/2, 2$ 인 경우의 오류 분배율 함수를 이용하였다.

γ 족은 다음과 같은 형태이다. 만일 $\gamma \neq 0$ 이면

$$\alpha^*(t, \gamma) = \frac{\alpha(1 - e^{-\gamma \min(1,t)})}{(1 - e^{-\gamma})} \tag{4.2}$$

$$\beta^*(t, \gamma) = \frac{\beta(1 - e^{-\gamma \min(1,t)})}{(1 - e^{-\gamma})}$$

이고, 만일 $\gamma = 0$ 이면

$$\alpha^*(t, \gamma) = \alpha \min(1, t) \tag{4.3}$$

$$\beta^*(t, \gamma) = \beta \min(1, t)$$

이다. Lan과 DeMets (1983)는 $\rho = 1$ 혹은 $\gamma = 0$ 의 형태로 나타낼 수 있으며, Chang 등 (1998)은 $\gamma = -4$ 인 경우의 형태이다. Selwyn과 Fish (2004)는 $\rho = 3$ 일 때의 오류분배율 함수를 사용하였다.

오류 분배율 함수는 제 1종 오류 분배율 함수와 제 2종 오류 분배율 함수로 나누어 생각할 수 있다. 그 중 제 2종 오류 분배율 함수는 제 1종 오류 분배율 함수의 특성과 같게 사용하거나 서로 다른 특성을 갖는 오류 분배율 함수를 사용 할 수 있다. 여기서는 두 오류 분배율 함수가 같은 특성을 갖도록 하였다.

본 논문에서는 오류 분배율 함수로 ρ 족과 γ 족을 포괄할 수 있는 누적베타분포함수(Beta(a, b))를 사용한다. 누적베타분포함수는 $[0, 1] \rightarrow [0, 1]$ 인 비감소함수이기 때문에 오류 분배율 함수의 성질을 만족하며, 위치 및 척도모수를 조정하면 기존의 ρ 족과 γ 족을 포함하는 다양한 형태의 오류 분배율 함수를 만들 수 있는 장점이 있다.

기존의 제 1종 오류 분배율 함수의 형태는 식 (4.4)와 같다.

$$\text{DeMet 1 : } 2 - 2\Phi\left(\frac{z_{\alpha}}{\sqrt{t}}\right) \tag{4.4}$$

$$\text{DeMet 2 : } \alpha \log[1 + (e - 1)t]$$

$$\text{DeMet 3 : } \alpha t$$

$$\text{KM 1 : } \alpha t^{\frac{3}{2}}$$

$$\text{KM 2 : } \alpha t^2$$

$$\text{Chang (1998) : } \alpha \frac{1 - e^{4t}}{1 - e^4}$$

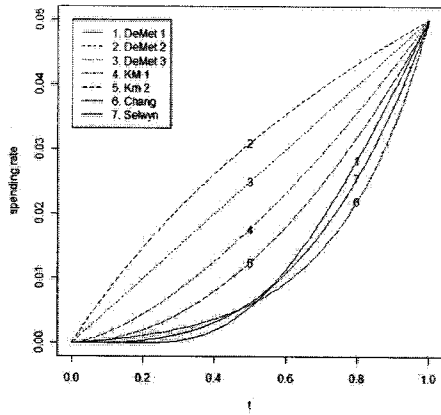
$$\text{Selwyn : } \alpha t^3$$

본 논문에서 제시하는 누적베타분포함수를 적용한 오류 분배율 함수는 식 (4.5)와 같다.

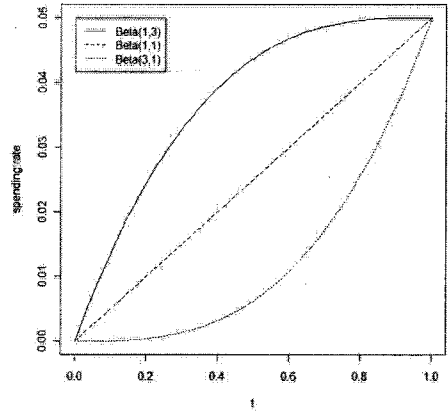
$$\alpha^*(t; a, b) = \alpha \frac{B_t(a, b)}{B(a, b)} \tag{4.5}$$

$$\beta^*(t; a, b) = \beta \frac{B_t(a, b)}{B(a, b)},$$

여기서, $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$, $B_t(a, b) = \int_0^t u^{a-1}(1 - u)^{b-1} du$, $0 < u < 1$ 이다. 본 논문에서는 오류분배율 함수의 특성에 따라 임계값이 어떻게 달라지는지를 보여주기 위해 Beta(1, 3), Beta(1, 1), Beta(3, 1) 3가지 함수를 이용한다.



(a) 기존의 제 1종 오류 분배율 함수



(b) 제시된 제 1종 오류 분배율 함수

그림 2: $\alpha = 0.05$ 인 경우 오류 분배율 함수의 비교

표 1: $\alpha = 0.05, \beta = 0.1, K = 2$ 인 경우 α, β 분배율 함수를 적용한 집단축차 검정의 임계값 및 성능 비교

N	n_1	n_2	a_1	a_2	r_1	r_2	α^*	$1 - \beta^*$	ASN_0	ASN_1	ASN^*	
$p_0 = 0.05, p_1 = 0.20$												
Beta(1, 3)	40	20	20	1	4	4	5	0.046	0.888	25.0	26.8	25.9
Beta(1, 1)	40	20	20	0	5	4	6	0.024	0.854	32.5	28.0	30.2
Beta(3, 1)	40	20	20	0	4	5	5	0.047	0.920	32.8	32.4	32.6
Fleming	40	20	20	0	4	4	5	0.052	0.922	32.5	28.0	30.3
Chang	40	20	20	0	4	5	5	0.047	0.920	32.8	32.4	32.6
$p_0 = 0.10, p_1 = 0.30$												
Beta(1, 3)	50	25	25	3	9	6	10	0.044	0.946	30.1	29.0	29.6
Beta(1, 1)	50	25	25	3	8	7	9	0.049	0.957	30.7	32.7	31.7
Beta(3, 1)	50	25	25	2	9	8	10	0.024	0.956	36.5	37.6	37.0
Fleming	50	25	25	2	8	7	9	0.057	0.976	36.3	33.3	34.8
Chang	50	25	25	4	8	6	9	0.049	0.906	26.6	27.6	27.1
$p_0 = 0.20, p_1 = 0.40$												
Beta(1, 3)	55	30	25	7	16	11	17	0.046	0.922	35.3	36.2	35.8
Beta(1, 1)	55	30	25	7	16	11	17	0.046	0.922	35.3	36.2	35.8
Beta(3, 1)	55	30	25	5	16	13	17	0.037	0.936	44.2	44.3	44.2
Fleming	55	30	25	6	16	12	17	0.039	0.933	39.6	40.3	40.0
Chang	55	30	25	7	16	11	17	0.046	0.922	35.3	36.2	35.8
$p_0 = 0.30, p_1 = 0.50$												
Beta(1, 3)	60	30	30	10	25	14	26	0.049	0.885	36.9	37.3	37.1
Beta(1, 1)	60	30	30	10	24	15	25	0.041	0.903	37.6	41.4	39.5
Beta(3, 1)	60	30	30	8	24	17	25	0.036	0.921	47.0	51.0	49.0
Fleming	60	30	30	9	24	16	25	0.037	0.916	42.1	46.5	44.3
Chang	60	30	30	10	24	15	25	0.041	0.903	37.6	41.4	39.5

그림 2를 살펴보면 Beta(1, 3) 함수는 초기 검정에서 오류율을 상당히 많이 사용한다는 것을 알 수 있다. 이는 Beta(1, 1) 함수에 비해 오목(concave)한 형태이며 반응변수가 연속형인 경우에 이 오류 분배율 함수를 이용하면 덜 보수적(conservative)인 검정이 된다. Beta(3, 1) 함수는 Beta(1, 1)에 비해 볼록(convex)한 형태이며, 검정초기에는 오류를 적게 사용하고 갈수록 오류를 많이 사용하는 형태를 지닌다. 검정통계량이 이항분포를 따를 때 최대 검정 횟수와 모수 p , 총 표본 크기, 각 검정 단계에서의

표 2: $\alpha = 0.05, \beta = 0.1, K = 3$ 인 경우 α, β 분배율 함수를 적용한 집단축차 검정의 임계값 및 성능 비교

	N	n_1	n_2	n_3	a_1	a_2	a_3	r_1	r_2	r_3	α^*	$1 - \beta^*$	ASN_0	ASN_1	ASN^*
$p_0 = 0.05, p_1 = 0.20$															
Beta(1, 3)	40	15	15	10	0	2	5	3	5	6	0.047	0.861	23.8	21.9	22.9
Beta(1, 1)	40	15	15	10	0	2	4	4	5	5	0.044	0.896	24.5	26.1	25.3
Beta(3, 1)	40	15	15	10	0	1	4	5	5	5	0.045	0.903	26.6	29.2	27.9
Fleming	40	15	15	10	-1	2	4	4	5	5	0.046	0.913	31.6	26.8	29.2
Chang	40	15	15	10	0	1	4	4	5	5	0.046	0.904	26.5	26.3	26.4
$p_0 = 0.10, p_1 = 0.30$															
Beta(1, 3)	50	20	20	10	2	6	9	6	8	10	0.045	0.942	26.7	27.9	27.3
Beta(1, 1)	50	20	20	10	2	7	8	6	9	9	0.035	0.926	26.4	28.0	27.2
Beta(3, 1)	50	20	20	10	0	6	8	7	9	9	0.048	0.970	38.3	33.0	35.6
Fleming	50	20	20	10	0	6	8	6	8	9	0.060	0.972	37.9	28.6	33.3
Chang	50	20	20	10	2	7	11	5	9	12	0.049	0.914	25.7	24.3	25.0
$p_0 = 0.20, p_1 = 0.40$															
Beta(1, 3)	55	20	20	15	4	11	17	8	14	18	0.047	0.878	27.5	28.7	28.1
Beta(1, 1)	55	20	20	15	3	10	16	9	14	17	0.043	0.925	33.5	34.0	33.8
Beta(3, 1)	55	20	20	15	2	10	16	11	15	17	0.035	0.927	38.1	41.6	39.9
Fleming	55	20	20	15	2	10	16	10	14	17	0.040	0.930	37.9	37.6	37.8
Chang	55	20	20	15	4	10	16	9	14	17	0.040	0.903	28.9	33.0	30.9
$p_0 = 0.30, p_1 = 0.50$															
Beta(1, 3)	60	20	20	20	6	14	25	11	18	26	0.046	0.871	29.9	33.3	31.6
Beta(1, 1)	60	20	20	20	5	14	24	12	19	25	0.039	0.906	34.8	39.7	37.2
Beta(3, 1)	60	20	20	20	3	13	24	13	20	25	0.037	0.920	43.6	45.6	44.6
Fleming	60	20	20	20	4	14	24	13	19	25	0.038	0.912	38.7	42.7	40.7
Chang	60	20	20	20	5	14	24	11	19	25	0.047	0.909	34.4	36.1	35.3

표본 크기 변화에 따라 채택점과 기각점, 제 1종 오류율, 검정력, ASN을 비교해 보면 여러 가지 특징이 나타난다. 우선 초기 검정에서 오류를 많이 분배하는 오류 분배율 함수를 사용하면 ASN은 낮아지는 경향이 있다. 두 번째로 오류분배율 함수의 형태가 틀림에도 불구하고 같은 기각역을 가질 수 있다. 이는 이산형의 성질과 수준 α 검정의 특성에 기인한다. 따라서 검정통계량이 이항분포를 따를 경우에는 Beta(1, 3), Beta(1, 1), Beta(3, 1) 3개의 경우만 살펴봐도 그 특성을 알아보기에 충분하다.

표 1, 2, 3은 $\alpha = 0.05, \beta = 0.1$ 인 경우, p_0, p_1 의 조건에 따라 각각 $K = 2, 3, 4$ 인 경우 기각점과 채택점을 나타낸다. 또한 오류 분배율 함수를 적용한 집단축차 검정법의 성능을 알아보기 위해 Fleming (1982)과 Chang 등 (1987)의 방법과 비교한다. 이때 ASN_0 는 귀무가설 하에서의 ASN, ASN_1 는 대립가설 하에서의 ASN 그리고 ASN^* 는 ASN_0 와 ASN_1 의 평균을 의미한다.

제 1종 오류를 0.05로 고정시킨 상태에서 제 2종 오류를 0.1보다 크게 준 경우에 제 1종, 제 2종 오류 분배율 함수를 적용한 집단축차 검정법으로 구한 채택점과 기각점에서는 주어진 제 1종 오류율과 제 2종 오류율을 넘지 않기 때문에 제 1종 오류는 0.05, 제 2종 오류는 0.1인 경우만 비교한다.

$\alpha = 0.05, \beta = 0.1$ 그리고 $K = 2$ 인 표 1을 살펴보자. $p_0 = 0.05, p_1 = 0.20$ 에서 Beta(1,3)인 경우 제 1종 오류는 0.046, 검정력은 0.888, ASN^* 는 25.9가 되며, 검정력의 경우 원하는 수준에 조금 못 미친다는 것을 알 수 있다. 반면 Beta(3,1)인 경우 제 1종 오류는 0.047, 검정력은 0.920, ASN^* 는 32.6으로 Chang의 결과와 같다. Fleming의 경우는 제 1종 오류가 0.052로 주어진 제 1종 오류율보다 높게 나타난다.

$p_0 = 0.30, p_1 = 0.50$ 인 경우를 살펴보면 Beta(1,1)인 경우 제 1종 오류는 0.041, 검정력은 0.903, ASN^* 는 39.5로 Chang의 결과와 같다. 반면 Beta(3,1)인 경우 제 1종 오류는 0.036, 검정력은 0.921, ASN^* 는 49.0으로 Beta(1, 1)인 경우보다 제 1종 오류율은 작아진 반면, 검정력은 높아지고 ASN^* 는 커

표 3: $\alpha = 0.045, \beta = 0.1, K = 4$ 인 경우 α, β 분배율 함수를 적용한 집단축차 검정의 임계값 및 성능 비교

	N	n_1	n_2	n_3	n_4	a_1	a_2	a_3	a_4	r_1	r_2	r_3	r_4	α^*	$1 - \beta^*$	ASN ₀	ASN ₁	ASN*
$p_0 = 0.05, p_1 = 0.20$																		
Beta(1,3)	40	10	10	10	10	0	0	0	4	3	4	5	5	0.048	0.848	21.4	20.2	20.8
Beta(1,1)	40	10	10	10	10	0	0	0	4	3	4	5	5	0.048	0.848	21.4	20.2	20.8
Beta(3,1)	40	10	10	10	10	0	0	0	4	5	5	5	5	0.039	0.844	21.9	25.7	23.8
Fleming	40	10	10	10	10	-2	0	2	4	4	4	5	5	0.049	0.912	26.9	25.1	26.0
Chang	40	10	10	10	10	-	-	-	-	-	-	-	-	-	-	-	-	-
$p_0 = 0.10, p_1 = 0.30$																		
Beta(1,3)	50	10	15	15	10	0	3	7	9	4	6	9	10	0.047	0.917	22.4	21.7	22.0
Beta(1,1)	50	10	15	15	10	0	2	7	8	5	7	9	9	0.033	0.923	25.8	27.4	26.6
Beta(3,1)	50	10	15	15	10	0	0	6	8	6	8	9	9	0.044	0.947	30.3	31.9	31.1
Fleming	50	10	15	15	10	-2	2	6	8	5	7	8	9	0.057	0.968	31.9	28.7	30.3
Chang	50	10	15	15	10	0	3	5	9	4	6	9	10	0.050	0.900	23.2	21.9	22.6
$p_0 = 0.20, p_1 = 0.40$																		
Beta(1,3)	55	15	15	15	10	2	7	12	17	7	11	15	18	0.049	0.894	27.0	27.6	27.3
Beta(1,1)	55	15	15	15	10	2	6	12	16	8	11	15	17	0.049	0.916	29.3	30.9	30.1
Beta(3,1)	55	15	15	15	10	0	5	12	16	9	13	15	17	0.040	0.930	38.7	38.0	38.4
Fleming	55	15	15	15	10	0	6	12	16	9	12	15	17	0.041	0.927	35.2	36.9	36.1
Chang	55	15	15	15	10	1	7	12	19	8	13	14	20	0.050	0.903	31.2	34.5	32.9
$p_0 = 0.30, p_1 = 0.50$																		
Beta(1,3)	60	15	15	15	15	3	10	17	25	9	15	20	26	0.047	0.876	29.6	31.6	30.6
Beta(1,1)	60	15	15	15	15	3	9	16	24	10	15	20	25	0.050	0.912	32.7	35.2	34.0
Beta(3,1)	60	15	15	15	15	1	8	16	24	11	17	21	25	0.037	0.916	40.1	43.0	41.5
Fleming	60	15	15	15	15	1	9	16	24	11	16	20	25	0.045	0.918	35.5	41.0	38.2
Chang	60	15	15	15	15	-	-	-	-	-	-	-	-	-	-	-	-	-

졌다.

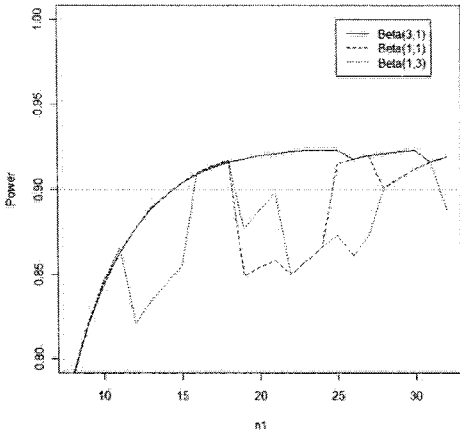
표 2는 $\alpha = 0.05, \beta = 0.1$ 그리고 $K = 3$ 인 경우의 결과를 나타낸다. $p_0 = 0.10, p_1 = 0.30$ 인 경우를 살펴보면 오류 분배율 함수를 적용한 집단축차 검정법은 주어진 제 1종 오류와 검정력을 모두 만족한다는 것을 알 수 있다. Chang의 결과와 비교해보면 검정력은 모두 높게 나왔으나 ASN*는 역시 조금 크다는 것을 알 수 있다.

$\alpha = 0.05, \beta = 0.1$ 그리고 $K = 4$ 인 표 3에서 $p_0 = 0.05, p_1 = 0.20$ 인 경우를 살펴보면 Chang은 조건에 맞는 임계값을 찾을 수가 없으며, α, β 오류 분배율 함수를 적용한 집단축차 검정법은 전부 검정력이 0.9에 못 미친다는 것을 알 수 있다.

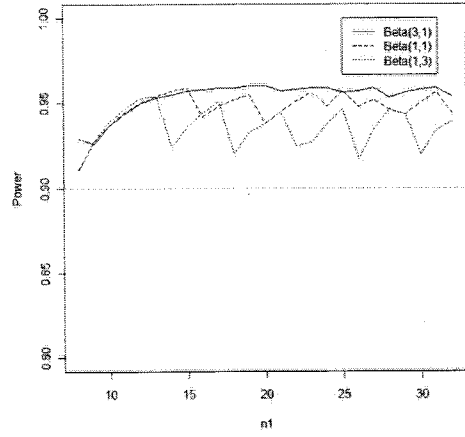
$p_0 = 0.30, p_1 = 0.50$ 인 경우 Chang은 조건에 맞는 임계값을 찾을 수가 없지만 Beta(1,1)과 Beta(3,1)의 경우 주어진 제 1종 오류와 검정력을 모두 만족한다. 이때 ASN*는 각각 34.0과 41.5이다.

오류 분배율 함수를 적용한 집단축차 검정법은 검정 초기에 오류율을 많이 사용하는 함수를 이용할 경우 검정력과 ASN이 낮아지는 경향이 있다. 집단축차 검정은 주어진 제 1종, 제 2종 오류율을 만족하면서 ASN을 최소화하는 임계값을 찾으므로 단순히 ASN을 낮추기 위해서 오목한 함수를 이용하면 검정력을 만족시키지 못하는 경우가 발생하기도 한다.

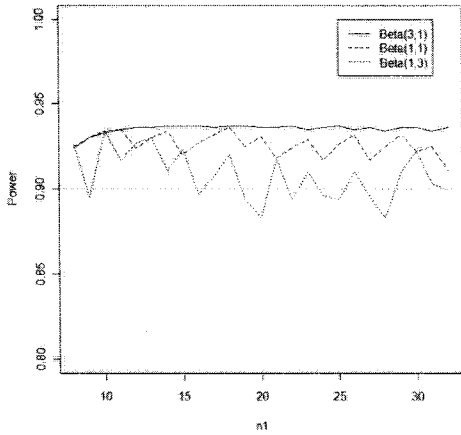
그림 3은 $\alpha = 0.05, \beta = 0.10, K = 2$ 일 때 검정력을 비교한 그림이다. 여기서 X-축은 첫 번째 단계의 표본크기를 나타낸다. 전체적으로 Beta(1,1)과 Beta(3,1)인 경우는 첫 번째 검정의 표본크기에 상관 없이 검정력이 0.9를 상회한다. 즉, $K = 2$ 인 경우 볼록한 함수를 이용하면 첫 번째 정보시점 t_1 에 상관 없이 항상 검정력이 0.9를 상회하는 결과를 얻을 수 있다. 하지만 검정력을 높이기 위해서 상대적으로 볼록한 함수를 이용하면 그만큼 ASN도 높아지기 때문에 적당히 조절하는 것이 필요하다. Beta(1,3)인



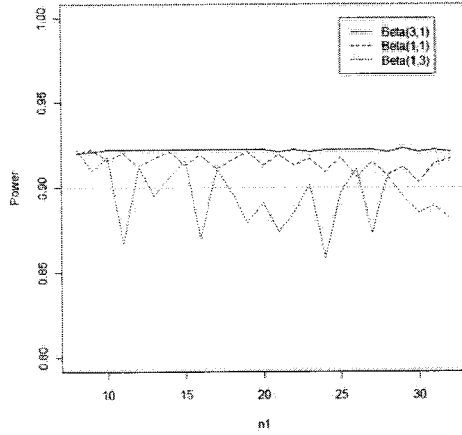
(a) $p_0 = 0.05, p_1 = 0.20, N = 40$



(b) $p_0 = 0.10, p_1 = 0.30, N = 50$



(c) $p_0 = 0.20, p_1 = 0.40, N = 55$



(d) $p_0 = 0.30, p_1 = 0.50, N = 60$

그림 3: $\alpha = 0.05, \beta = 0.1, K = 2$ 일 때 p_0 와 p_1 에 따른 검정력 비교

경우는 검정력이 0.9를 만족하지 못하는 경우가 발생한다. 즉, 검정력이 0.9를 만족하기 위해서 정보시점의 조절이 필요하다.

이상과 같이 이산형 반응변수인 경우 오류 분배율 함수를 적용한 집단축차 검정법은 오류 분배율 함수의 형태, 정보 시점에 따라 원하는 검정력을 얻지 못하는 경우가 생기기도 한다. 특히 오목한 형태의 오류 분배율 함수를 이용하는 경우 이러한 현상이 두드러지게 나타난다. 반면, 볼록한 형태의 오류 분배율 함수를 이용하는 경우 오목한 형태의 오류 분배율 함수를 이용하는 경우보다 검정력이 상대적으로 높게 나타나며, 더불어 ASN도 높게 나타난다. 따라서 임상시험의 성격에 따라 적절한 정보시점과 오류 분배율 함수를 이용하면 최적의 임계점을 찾을 수 있다.

오류 분배율 함수를 적용한 집단축차 검정법은 여러 가지 장점을 지닌다. 첫 번째는 유연한 검정이 가능하다는 것이다. 기존의 집단축차 검정법은 임상시험 시행전에 검정 시점이 정해져 있다. 이는 원활한 임상시험이 진행되지 않을 시, 검정이 용이하지 않을 수 있음을 의미한다. 하지만 오류 분배율 함

수를 적용한 집단축차 검정은 임상시험 시행전에 오류 분배율 함수만 결정하면 되므로 검정 시점에 제한이 없다. 실질적으로 임상시험은 계획된 대로 시험이 진행되지 않는 경우에는 검정시점의 변경이 불가피하게 되고, 이런 경우에 오류 분배율 함수를 적용한 집단축차 검정은 아무런 제약이 없다. 두 번째로 임계값을 찾기 위한 연산시간이 짧다. Fleming (1982)의 방법은 연산시간이 짧기는 하나 주어진 제 1종 오류를 만족하지 못하는 경우가 빈번히 발생한다. Chang 등 (1987)과 Simoń (1988)의 경우 기본적으로 모든 조합을 고려하기 때문에 연산시간이 오래 걸리는 단점이 있다. 특히 총 검정 횟수가 많을수록 이러한 현상은 두드러지게 나타난다. 하지만 오류 분배율 함수를 적용한 집단축차 검정은 총 검정수에 관계없이 연산시간이 짧다는 장점이 있다. 예를 들어 $(p_0, p_1) = (0.3, 0.5)$, $\alpha = 0.05$, $\beta = 0.1$, $N = 60$ 인 조건에서 오류 분배율 함수를 적용하지 않은 집단축차 검정법인 Chang 등 (1987)의 방법은 모든 가능한 경우를 고려하여 임계값이 결정되므로 $K = 2$ 인 경우에 비해 $K = 3$ 인 경우는 19배, $K = 4$ 인 경우는 2,181배의 계산 시간이 소요된다. 반면 오류 분배율 함수를 적용한 집단축차 검정은 $K = 2, 3, 4$ 인 경우 거의 동일한 계산 시간이 소요된다.

5. 결론

본 연구에서는 이산형, 특히 이항 반응변수인 경우 제 1종 오류 분배율 함수와 제 2종 오류 분배율 함수를 적용한 집단축차 검정법을 제안하였다. 이는 제 1종 오류 분배율 함수만 적용한 방법으로 쉽게 적용이 가능하며, 또한 제 2종 오류 분배율 함수만 적용한 방법으로도 적용이 가능하다.

반응변수가 연속형인 경우 오류 분배율 함수를 적용한 집단축차 검정은 주어진 오류율의 손실없이 임계값을 찾는 것이 가능하다. 즉, 주어진 제 1종 오류와 제 2종 오류를 원하는 만큼 분배하는 임계값을 찾을 수 있다. 하지만 반응변수가 이산형인 경우 주어진 제 1종 오류율을 만족하는 크기 α 검정이 불가능하다. 따라서 본 연구에서는 크기 α 검정 대신 수준 α 검정을 이용하였으며, 각 검정 단계에서 사용되어야 할 오류율을 이용하는 대신, 각 검정단계까지 사용되어야 할 누적 오류율을 이용하였다. 그 결과 각 검정단계까지 사용되어야 할 누적 오류 분배율 함수를 적용한 경우 주어진 제 1종 오류율에 근접하면서도 높은 검정력을 보였으며, ASN은 상대적으로 낮게 나왔다.

본 연구에서는 누적베타분포함수를 오류 분배율 함수로 이용하였다. 누적베타분포는 위치모수와 척도모수의 조정으로 기존에 연구된 오류 분배율 함수를 포함할 수 있는 장점이 있다. 특히 오목한 함수 형태와 볼록한 함수 형태를 다양하게 표현 할 수 있어 오류 분배율 함수를 원하는 형태로 쉽게 표현이 가능하다.

오류 분배율 함수를 적용한 방법과 기존의 집단축차 검정법을 비교 하였을 때, 요구하는 검정력이 상당히 높지 않은 경우는 기존의 집단축차 검정법에 비해 ASN이 작으면서도 검정력의 손실이 미비하였다. 또한 제 1종 오류를 원하는 수준에 근접시키기 때문에 중간분석의 목적에 더욱 부합됨을 알 수 있다. 요구하는 검정력이 상당히 높은 경우에 기존의 집단축차 검정법은 적당한 임계값을 제시하지 못하는 경우가 많다. 하지만 오류 분배율 함수를 적용한 방법은 볼록한 형태의 오류 분배율 함수를 적용함으로써 원하는 수준의 오류율을 만족하는 임계값을 찾을 수 있다.

기존의 집단축차 검정법은 제약이 많은 반면에 오류 분배율 함수를 적용한 집단축차 검정법은 상대적으로 제약이 적다. 그만큼 유연한 검정이 가능하게 되며, 임상시험이 계획대로 원활하게 진행되지 못하는 경우에도 유연하게 대처할 수 있는 장점이 있다. 또한 기존의 집단축차 검정법은 총 검정 단계수가 커질수록 연산시간이 매우 길어진다는 단점이 있으나 오류 분배율 함수를 적용한 집단축차 검정법은 연산시간이 총 검정 단계수에 크게 구애받지 않는다.

결론적으로 이항 반응변수인 경우 α, β 오류 분배율 함수를 적용한 집단축차 검정법은 적은 제약에 비해 높은 성능을 지니며, 임상시험의 성격에 따라 유연한 검정이 가능하다.

참고 문헌

- Anderson, K. M. (2006). Optimal spending functions for asymmetric group sequential designs, *Biometrical Journal*, **48**, 1–9.
- Armitage, P., McPherson, C. K. and Rowe, B. C. (1969). Repeated significance test on accumulating data, *Journal of the Royal Statistical Society, Series A*, **132**, 235–244.
- Barber, S. and Jennison, C. (2002). Optimal asymmetric one-sided group sequential tests, *Biometrika*, **89**, 49–60.
- Chang, M. N., Therneau, T. M., Wieand, H. S. and Cha, S. S. (1987). Designs for group sequential phase II clinical trials, *Biometrics*, **43**, 865–874.
- Chang, M. N., Hwang, I. K. and Shih, W. J. (1998). Group sequential designs using both type I and type II error probability spending functions, *Communications in Statistics - Theory and Methods*, **27**, 1323–1339.
- Fleming, T. R. (1982). One-sample multiple testing procedure for phase II clinical trials, *Biometrika*, **38**, 143–151.
- Hwang, I. K., Shih, W. J. and DeCani, J. S. (1990). Group sequential designs using a family of type I error probability spending functions, *Statistics in Medicine*, **9**, 1439–1445.
- Kim, K. and DeMets, D. L. (1987). Design and analysis of group sequential tests based on the type I error spending rate function, *Biometrika*, **74**, 149–154.
- Lan, K. K. G. and DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials, *Biometrika*, **70**, 659–663.
- Lan, K. K. G. and DeMets, D. L. (1984). More flexible sequential and non sequential designs in long-term clinical trials, *Communications in Statistics - Theory and Methods*, **13**, 2339–2353.
- O'Brien, P. C. and Fleming, T. R. (1979). A multiple testing procedure for clinical trials, *Biometrics*, **35**, 549–556.
- Pocock, S. J. (1977). Group sequential in design and analysis of clinical trials, *Biometrika*, **64**, 191–199.
- Schultz, J. R., Nichol, F. R., Elfring, G. L. and Weed, S. D. (1973). Multiple stage procedures for drug screening, *Biometrics*, **29**, 293–300.
- Selwyn, M. R. and Fish, S. M. (2004). Choice of alpha spending function and time points in clinical trials with one or two interim analysis, *Pharmaceutical Statistics*, **3**, 193–203.
- Simon, R. (1988). Optimal two-stage designs for phase II clinical trials, *Controlled Clinical Trials*, **10**, 1–10.

Group Sequential Tests Using both Type I and Type II Error Spending Rate Functions on Binomial Response

Donguk Kim^{1,a}, Jin Hyun Nam^a

^aDepartment of Statistics, Sungkyunkwan University

Abstract

In interim analysis, group sequential tests are widely used for the ethical, scientific, and economic reasons. In this paper, we propose the group sequential tests using both type I and type II error spending rate functions when the response variable is discrete, especially binomial distribution, in the interim analysis. In addition, we propose new error spending rate function which covers the formerly proposed. Our method has good property that is flexible, fast and easily applicable. A numerical simulations are carried out to evaluate our method and it shows good performance.

Keywords: Discrete variable, error spending rate function, group sequential, interim analysis.

¹ Corresponding author: Professor, Department of Statistics, Sungkyunkwan University, 53 Myeongnyun-Dong 3-Ga, Jongno-Gu, Seoul 110-745, Korea. E-mail: dkim@skku.edu