# 다중 검정에서의 효과적인 거짓 발견 확률 제어 방법 연구

전유경
성신여자대학교 수리통계데이터사이언스학부 통계학전공

Summary
다중 검정을 보정 할 때, 오류율이 총 가설의 개수에 의존하는 FWER(Family Wise Error Rate)과 달리 귀무가설을 기각시킨 결정 중에서 잘못 기각한 가설의 개수를 살펴보는 FDR(False Discovery Rate)이 있다. FDR을 제어하는 방법인 Benjamini-Hochberg procedure는 많은 분야의 다중 검정 문제에서 제1종 오류를 제어하면서 검정력이 낮아지는 것을 막는다. 하지만 Benjamini-Hochberg procedure는 검정하는 가설들이 서로 독립일 때에만 적용이 가능한데, 실제로 많은 문제에서는 독립적이지 않은 상황에서 검정을 시행해야 한다. 그래서 이 FDR 제어 방법을 확장하여 양의 의존성 조건을 가지는 다중 검정에서도 제1종 오류를 제어할 수 있도록 한다. 그리고 제안된 모든 다중 검정 보정 방법들을 비교하여 어떤 방법이 가장 효과적인지 살펴본다.

## 1. Introduction

통계적 가설 검정은 모집단의 실제 값에 대하여 표본의 정보를 이용하여 가설의 합당성 여부를 판정하는 과정을 의미한다. 이때 다중 검정이란 여러 개의 가설 검정을 동시에 수행하는 것이다. 예를 들어 성별, 학업 여부, 일의 형태에 따른 월급 평균 차이를 분석하고자 할 때, 각 조건마다 해당하는 가설을 동시에 검정하여 월급 평균 차이를 분석할 수가 있다. 이러한 다중 검정에서 동시에 여러 개의 가설로 검정을 수행한다는 사실을 고려하지 않고 진행한다면, 많은 제1종 오류(false positive)가 발생하게 될 것이다. 자세히 설명하자면, 단일 검정에서 p-value가 0.05보다 작을 때 귀무가설을 기각한다고 하면 기각된 가설이 제1종 오류일 가능성이 5%임을 의미한다. 이를 m개의 가설로 다중 검정을 시행한다면, m개의 가설에 대해서 제1종 오류일 가능성이 0.05 x m 이 될 것이다. 즉 더 높은 유의 수준을 적용한 것처럼 되어, 더 많은 가설에 대해 잘못된 기각을 내릴 확률이 증가하게 된다. 이러한 문제는 실제로 참인 가설을 잘못 판단하는 위험을 높이게 되어 연구 결과의 신뢰성을 낮춘다. 그래서 이와 같은 위험을 제어하는 방법을 찾는 것은 다중 검정에서 매우 중요한 과정이다.

제1종 오류의 위험을 제어하는 방법으로 먼저 제시된 것은 바로 FWER(Family Wise Error Rate)이다. FWER이란 다수의 가설을 검정할 때, 적어도 하나의 가설이 제1종 오류를 낼 확률이다. 이러한 FWER을 제어하여, 다중 검정에서 제1종 오류의 확률이 누적으로 증가하게 되는 것을 막는 것이다. 제어 방법으로는 'Bonferroni Method'와 'Holm's Method'가 있다.

먼저 Bonferroni Method는 각각의 가설 검정의 유의 수준을 총 가설의 개수로 나눈 값으로 유의 수준을 조정하여 전체 오류율 FWER이 유의 수준 $\alpha$ 이하가 되도록 통제하는 방법이다. 하지만 Bonferroni Method는 총 가설의 개수가 많아지면 개별 검정의 유의 수준이 작아지기 때문에 귀무가설을 잘 기각시키지 않아 매우 보수적이게 된다. 이는 귀무가설이 거짓일 때 해당 귀무가설을 기각시키는 true positive를 잡지 않아 귀무가설이 거짓일 때 해당 귀무가설을 기각시키지 않는 false negative를 발생시켜, 제2종 오류($\beta$)가 증가하게 된다. $\beta$의 증가는 곧 귀무가설이 거짓일 때 귀무가설을 기각할 확률인 검정력(1-$\beta$)이 낮아진다는 것이다.

이러한 Bonferroni Method의 문제점을 보완하고자 만들어진 것이 바로 Holm's Method이다. Holm's Method는 총 가설의 개수가 m개인 다중 검정에서 개별 검정의 p-value를 1부터 m까지 오름차순으로 나열한 뒤 유의수준 $\frac{\alpha}{(m+1-i)}$ (i = 1, ..., m)을 각각 다른 p-value cutoff로 부여하는 방법이다. 이 방법에서

제시한 기각역은 Bonferroni Method의 기각역보다 크기 때문에 덜 보수적인 방법이라고 할 수 있다. 하지만 여전히 FWER 제어 방법이 총 가설의 개수에 의존하고 있어 검정력이 낮아지는 문제는 피할 수 없다. 또한 연구자들은 제1종 오류 가능성을 통제하는 것보다 개별 가설에 대한 검정력을 유지하는 것을 더 선호하기에 많은 연구에서 FWER 통제가 간과되는 경우가 많다.

이에 본 연구에서는 다른 오류율로 다중 검정의 문제점을 해결하고자 하였고, 다중 검정에서의 각 가설들이 독립적인 상황만이 아닌 의존적일 때에도 제1종 오류를 통제할 수 있는 방법을 함께 연구하였다.

본 연구의 진행은 다음과 같다. 먼저 2절에서는 FDR(False Discovery Rate)이 무엇인지 설명하고 FWER과의 관계를 보인다. 또한 해당 오류율을 제어하는 방법을 제시하고, 검정하는 각 가설들이 의존적일 때에는 어떻게 제어하는지 소개한다. 그리고 3절에서는 1, 2절에서 소개한 방법들의 기각 수와 검정력에 대한 시뮬레이션 연구를 보여준다. 4절에서는 실제 자료를 이용한 분석 결과를 제시하여 제안된 방법이 잘 적용됨을 실험을 통해 보일 것이다. 마지막 5절에서는 결론으로 본 연구 결과에 대해 정리한다.

## 2. False Discovery Rate

| | $H_0$ is True | $H_0$ is False | Total |
|---|---|---|---|
| Reject $H_0$ | $V$ | $S$ | $R$ |
| Do Not Reject $H_0$ | $U$ | $W$ | $m-R$ |
| Total | $m_0$ | $m-m_0$ | $m$ |

V: 제 1종 오류인 가설의 수 S: true positive인 가설의 수

U: true negative인 가설의 수 W: 제 2종 오류인 가설의 수

Table1. m개의 귀무가설을 검정할 때 발생한 오류의 수

FDR(False Discovery Rate)이란 귀무가설을 기각시킨 결정 중에서 잘못 기각한 경우(false positive)에 대한 기대 비율이며, 어느 정도의 false positive를 수용하면서 더 많은 true positive를 찾는 것을 목표로 한다.

$$\text{FDR} = E(Q) = E\{\frac{V}{(V+S)}\} = E(\frac{V}{R}) \quad (R>0)$$

여기에서 기댓값을 씌우는 이유는 실제로 귀무가설이 참인지 거짓인지 알 수 없기에 기댓값을 씌워 분석자가 수용할 수 있는 false positive의 수를 지정해주기 위함이다. 이러한 FDR은 비교하는 집단이 많고 진실인 귀무가설로만 구성되어 있지 않는 한, 여러 집단에서 어떤 수준에서 별도로 통제되면 전체 집단에서도 동일한 수준에서 통제할 수 있다. 만약 모든 귀무가설이 참이면 귀무가설을 기각하는 개수가 곧 false positive인 귀무가설의 수가 되므로 FDR과 FWER는 같게 된다. 또한 FDR은 FWER과 달리 많은 귀무가설이 기각된 경우에서 사용하는 것이 좋다. 왜냐하면 많은 귀무가설이 기각된다는 것은 상당수의 귀무가설이 실제로 거짓일 가능성이 높기에, 제1종 오류가 전체 결과에 미치는 영향은 상대적으로 낮게 된다. 즉, 단일 오류로 전체 결과가 왜곡될 확률이 적어 이 상황에선 오류의 절대 수 보단 그 비율을 제어하여 더 많은 유용한 발견(true positive)을 찾는 것이 중요하다.

### 2.1 Benjamini-Hochberg procedure

Benjamini-Hochberg procedure는 검정통계량이 독립적인 경우 FDR을 수준q에서 통제하는 절차이다. 절차 과정을 자세히 보면 다음과 같다. 총 가설의 개수가 m개인 다중 검정에서 개별 검정의 p-value를 1부터 m까지 오름차순으로 나열한 뒤 임계값 $\frac{i}{m}\times q$ (i = 1, ..., m) 보다 같거나 작은 p-value 중에서 가장 큰 i에

대하여 이 p-value와 모든 더 작은 p-value에 해당하는 가설을 기각하여 거짓 발견 확률을 수준q에서 통제하는 것이다. 이때 수준 q는 분석자가 정하는 것이고, 허용 가능한 최대 거짓 발견 비율이다. 따라서 Benjamini-Hochberg procedure는 원하는 수준 q에서 기각된 가설들의 예상 수를 최대화하여 FDR을 통제하여, 가능한 많은 유의미한 발견과 거짓 발견 비율을 낮추는 역할을 한다.

이러한 Benjamini-Hochberg procedure로 FWER도 통제할 수 있는데, 먼저 모든 가설이 참인 경우에 수준 q와 유의 수준 $\alpha$ 가 모두 1을 만족하게 되므로, FDR 제어 방법이 Simes' Global Test와 같아져 FWER을 통제할 수 있게 된다. 만약 참인 귀무 가설의 수가 전체 가설의 수보다 작다면 기존 FDR 제어 방법을 사용하여 FWER을 제어하기엔 다중 검정 보정 방법이 보수적이지 않다. 따라서 오름차순으로 정렬한 p-value의 임계값 $\frac{q}{(m+1-i)}$ (i = 1, ..., m)을 적용하면 더 엄격한 기준이 되므로, FWER을 효과적으로 통제할 수 있게 된다.

2.2 The Control of the FDR in multiple testing under dependency

Benjamini-Hochberg procedure는 검정통계량이 독립적인 경우에만 해당하지만, 주로 연구를 하게 되면 독립적인 상황보다는 검정통계량끼리 의존적인 상황이 더 많이 발생하곤 한다. 이런 상황에선 검정통계량 중에서 진실인 귀무가설 집합의 joint distribution이 PRDS를 만족하다면, Benjamini-Hochberg procedure는 수준 $\frac{m_0}{m}\times q$ ($m_0$: 진실인 귀무가설의 수, m: 총 가설의 수) 이하에서 FDR을 제어할 수 있게 된다. 이와 같이 FDR을 통제하는 것은 기존 q에서 $\frac{m_0}{m}$를 곱하는 것이라 거짓 양성 비율을 더 낮추게 된다. 그래서 원래의 Benjamini-Hochberg procedure보다 더 보수적이지만 더 정확한 검정이 가능해진다.

2.3 Positive Regression Dependency Subset

위 2.2 방법의 통제 조건에서 나온 PRDS(Positive Regression Dependency Subset)란 부분집합 $I_0$에 대하여 해당 분포가 양의 회귀 의존성을 갖는 성질, 즉 $I_0$에 속하는 특정 변수의 값이 증가함에 따라 나머지 변수들의 조건부 분포도 확률적으로 증가한다는 것이다. 이를 증가하는 집합 D에 적용하면, 증가하는 집합 D와 부분집합 $I_0$ 내의 각 인덱스 i (i=1, ..., n)에 대하여 확률 $P(X\in D|X_i=x)$가 $x$에서 비감소하는 성질이다. 즉 특정값 $x$를 증가시킬 때, X가 D에 속할 확률이 동일하거나 증가한다는 것이고, 이는 PRDS에서 특정 변수의 증가가 전체 시스템에 긍정적 결과를 가져온다는 것을 나타낸다.

이러한 PRDS는 두 가지 조건을 갖는데, 첫 번째는 단일 변수에 대한 조건부 단조성이다. 어떤 변수 $X_i$에 대해 한 변수가 증가(or 감소)할 때 다른 변수도 증가(or 감소)한다는 '단조성' 조건을 두고, 이 조건하에서 다른 변수를 살펴보는 것이다. 두 번째는 부분집합에 대한 조건부이다. PRDS는 전체 변수 집합이 아닌 선택한 부분집합에서만 양의 회귀 의존성을 확인해야 한다.

PRDS는 불변성이라는 특징을 갖는다. PRDS 정의를 포함하여 모든 속성들에 해당하는 각 변수에서, 여러 변수들 간의 상대적인 순서를 유지하면서 이들에 변수들이 서로 같은 방향으로 움직이도록 하는 변환을 적용하는 개념인 Comonotone Transformation을 하여도 불변한다는 것이다. 또한 검정통계량의 결합분포가 PRDS일 때, 이에 상응하는 p-value의 결합분포도 PRDS이다. 즉 각 실험에서 나온 검정통계량의 결합분포가 PRDS 속성을 가지면, 한 실험에서 유의미한 결과인 낮은 p-value를 얻었을 때 다른 실험에서도 유의미한 결과를 얻을 확률이 감소하지 않는다는 것이다.

또한 다변량 확률 분포를 PRDS에 접목시키면 다음과 같다. 집합 X가 $MTP_2$ 조건을 만족하면 어떤 부분집합이든 PRDS도 만족하게 된다. 여기서 $MTP_2$(Multivariate Totally Positive of order2)란 x와 y가 함께 증가(or 감소)하는 경우의 확률이 x or y 중에서 하나만 증가(or 감소)하는 경우의 확률보다 크거나 같다는

성질이다.

$$f(x) \cdot f(y) \le f(\min(x,y)) \cdot f(\max(x,y))$$

즉 두 변수 x,y는 강한 양의 의존성을 갖는다는 것이다.

PRDS에서 말하는 양의 회귀 의존성은 양의 연관성과 다르다. 양의 연관성이란 두 변수 사이에 양의 상관관계가 존재하는 성질이다. 즉 두 함수 f, g가 모두 증가(or 감소)하고 $cov(f(x)g(x)) \ge 0$ 을 만족하는 것이다. 겉보기엔 두 정의가 함축될 수 있다고 생각할 수 있지만, 엄연히 서로 다르다는 것을 두 가지 예시를 통해 보일 수 있다. 첫 번째 예시는 바로 상관계수이다. 양의 연관성에서는 상관계수가 음수가 아니면 양의 연관성은 존재한다고 볼 수 있는데, PRDS에서는 이 성질을 만족하기 위해서 모든 상관계수가 양수일 필요는 없다. 일부 음수인 상관계수가 있어도 전체적으로 PRDS를 만족할 수 있다. 두 번째 예시는 이변량 분포이다. 만약 이변량 정규분포에서 상관관계가 약하면, 양의 연관성을 만족할 순 있지만 양의 회귀 의존성은 이루어질 수 없을 것이다. 즉 양의 연관성이 있는 분포가 반드시 PRDS를 만족하지 않을 수 있음을 보인다.

하지만 PRDS를 함축하는 양의 연관성도 존재한다. 바로 조건적 양의 연관성이다. 이는 변수 집합 X를 두 부분 집합 $X_1$, $X_2$ 로 나눌 때, $X_2$가 $X_1$에 대한 함수 h($X_1$)에 조건을 둔다면 함수 h와 $X_2$가 양의 연관성을 갖는다는 것이다. 예를 들어 환자 집단 X에서 특정 증상을 $X_1$이라고 하고, 치료 반응을 $X_2$라고 라고 한다면, $X_1$의 특정 패턴인 함수 h($X_1$)이 주어지고 $X_2$가 해당 패턴과 양의 연관성을 가지면 이는 조건적 양의 연관성이라고 할 수 있다.

## 3. Simulation

지금까지 설명한 다중 검정 보정 방법들을 시뮬레이션을 통해 비교한다. 변수의 수가 200개이고 X1 집단과 X2 집단의 샘플 크기는 각각 50개이다. 이러한 시뮬레이션을 총 500번 반복한다. 이때 X1과 X2 집단은 독립적인 다변량 정규 분포에서 추출된 랜덤 변수이다. 먼저 독립적인 다변량 정규 분포이므로 두 집단의 공분산 행렬은 주 대각선의 원소가 1인 대각행렬로 나타낸다. 그리고 두 집단의 평균 벡터가 0으로 동일할 때를 귀무가설이라고 설정하고, X1은 귀무가설과 동일한 평균 벡터를 가지지만 X2는 절반은 1이고 나머지 절반은 0인 평균 벡터를 가질 때를 대립가설로 설정한다. 이렇게 설정한 시뮬레이션에서 살펴볼 것은 각 보정 방법별로 두 X1, X2 집단을 평균 비교하는 것이다. 이때 보정 방법은 1. No Correction, 2. Bonferroni Method, 3. Holm's Method, 4. Benjamini-Hochberg procedure을 이용하였다.

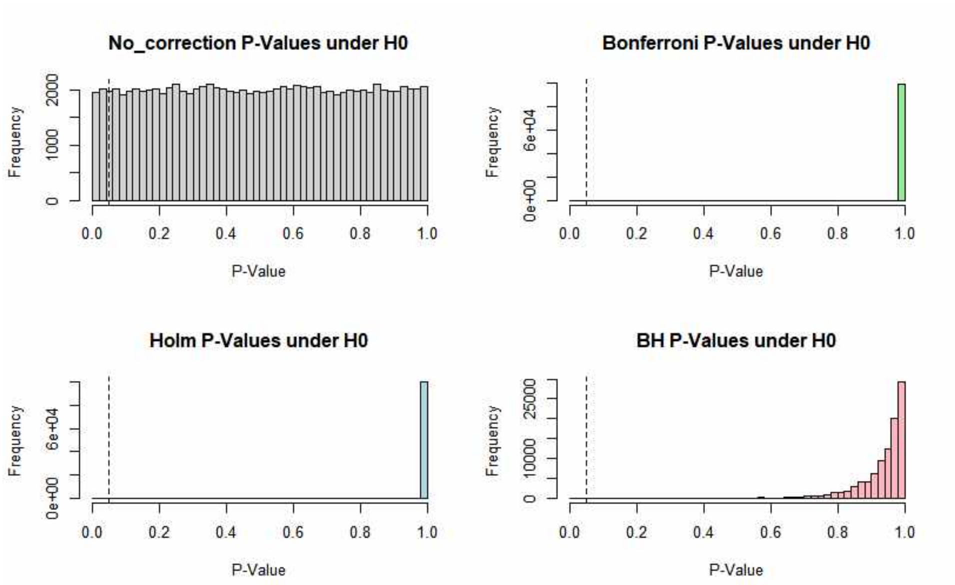


Figure1. 귀무가설 하에서 보정 방법별 p-value 분포 (점선: P-Value = 0.05)

귀무가설 하에서 살펴본 다중 검정 보정 방법별 p-value 분포인 Figure1을 살펴보면, 첫 번째로 No Correction에서 p-value가 균등 분포로 나타난다. 이는 귀무가설이 참일 때, 즉 실제로 X1과 X2가 차이가 없을 때 p-value가 0과 1 사이에서 균일하게 발생한다는 것을 나타낸다. 이때 4가지의 다중 검정 보정 방법에서 No Correction이 유독 귀무가설을 기각하는 수가 많다는 걸 알 수 있지만, 여전히 제1종 오류를 0.05로 유지하고 있다는 것을 알 수 있다. 두 번째로 FWER 제어 방법인 Bonferroni Method와 Holm's Method는 모든 p-value에 대하여 엄격한 보정을 수행하기 때문에 거의 모든 p-value가 1 근처에 있다는 것을 볼 수 있다. 세 번째로 FDR 제어 방법인 Benjamini-Hochberg procedure는 FWER 제어 방법들 보다는 덜 엄격하므로 p-value가 1 근처에만 분포되어 있진 않지만, 제1종 오류를 제어하고 있으므로 거의 모든 p-value가 0.05보다 큰 것을 볼 수 있다.

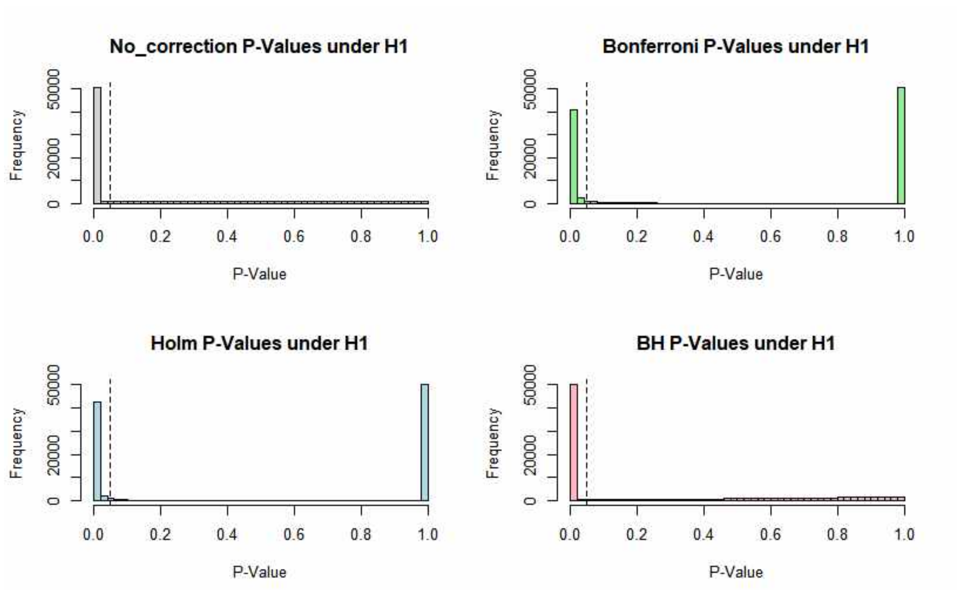

Figure2. 대립가설 하에서 보정 방법별 p-value 분포 (점선: P-Value = 0.05)

대립가설 하에서 살펴본 다중 검정 보정 방법별 p-value의 분포를 나타낸 Figure2를 살펴보면, 첫 번째로 No Correction에서 대부분의 p-value가 매우 낮은 구간에 집중되어 있음을 보인다. 이는 대립가설이 참일 때, 테스트가 실제 차이를 잘 감지하고 있음을 나타낸다. 두 번째로 FWER 제어 방법인 Bonferroni Method와 Holm's Method의 p-value는 0.05보다 작은 곳이나 1에 몰려 있음을 볼 수 있는데, 이는 제1종오류를 너무 강하게 제어하여 실제로 유의한 차이를 감지하는 능력인 검정력이 크게 감소했음을 유추해 볼 수 있다. 세 번째로 FDR 제어 방법인 Benjamini-Hochberg procedure는 FWER 제어 방법들 보다는 덜 보수적이므로 p-value가 1에 몰려 있지 않음을 알 수 있다. 또한 No Correction과 p-value 분포가 비슷하지만, false positive를 제어하기에 극단적으로 가설들을 기각하고 있지 않다는 것을 알 수 있다.

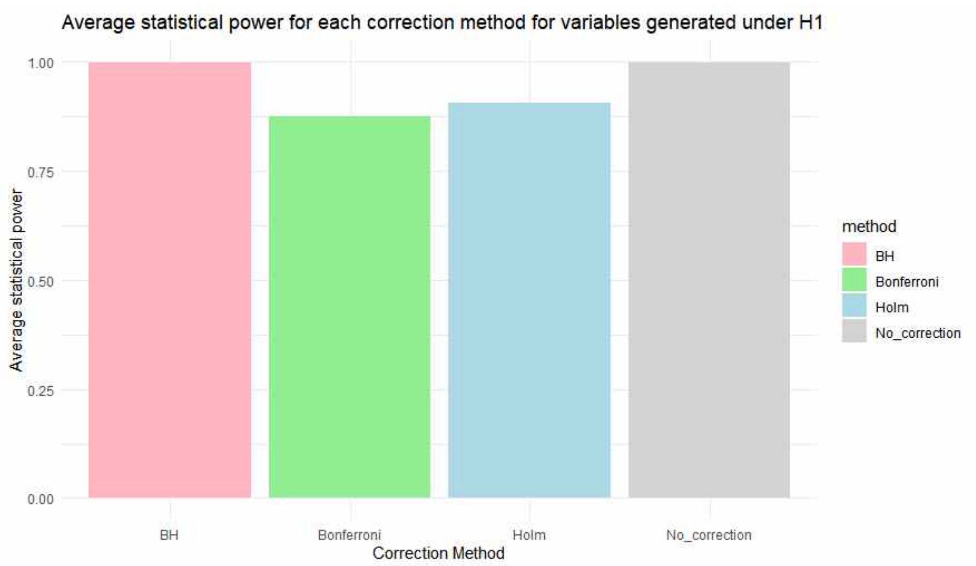


Figure3. 대립가설 하에서 생성된 변수에 대한 각 보정 방법별 평균 검정력

보정 방법별로 검정력을 비교하기 위해 Figure3을 살펴보면, No Correction이 가장 높은 검정력을 가진다는 것을 알 수 있고 Benjamini-Hochberg procedure, Holm's Method, Bonferroni Method 순으로 높은 검정력을 가진다는 것을 알 수 있다. 이는 앞의 1, 2장에서 FWER 제어 방법들 중 가장 엄격하게 제1종 오류를 제어하는 Bonferroni Method가 가장 낮은 검정력을 갖는 다는 것과 일치한다. Bonferroni Method에서 나타나는 낮은 검정력을 보완하고자 만들어진 Holm's Method가 Bonferroni Method 보다는 높은 검정력을 가지지만 여전히 No Correction만큼의 검정력을 가지지 못하는 것을 알 수 있다. 그래서 이를 보완하고자 제시한 FDR 제어 방법인 Benjamini-Hochberg procedure는 No Correction 만큼의 검정력을 나타내고 있다는 것을 알 수 있다.

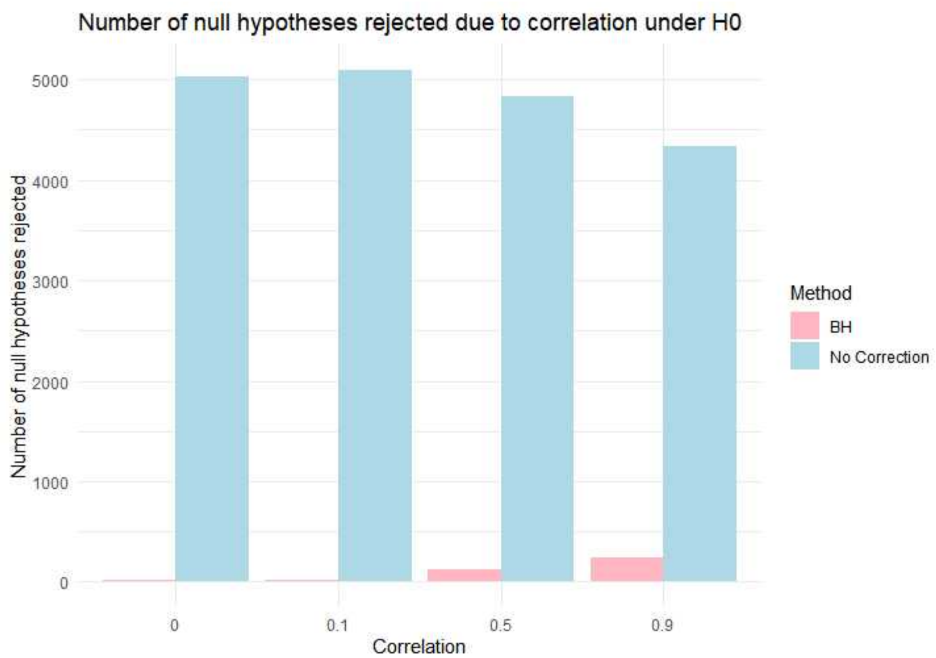


Figure4. 귀무가설 하에서 상관관계에 따른 귀무가설 기각 개수

다중 검정 시 검정하는 가설들 사이에 양의 의존성이 있을 때를 살펴보려고 한다. 이때의 시뮬레이션 조

건은 독립일 때의 시뮬레이션과 동일하게 진행하였다. Figure4는 귀무가설 하에서 상관관계가 0, 0.1, 0.5, 0.9 일 때 각각 다중 검정 보정 방법에 따른 귀무가설 기각 개수를 표현하였다. 여기서 Benjamini-Hochberg procedure는 제1종 오류를 제어하기에 No Correction 보다 훨씬 더 적게 귀무가설을 기각시킨다. 그리고 Benjamini-Hochberg procedure에서 상관관계가 증가할수록 기각하는 귀무가설의 개수도 증가하는 것으로 보이는데, 이는 비슷한 실험 조건이나 유사한 연구 가설로 인해 양의 상관관계를 가진다면, 하나의 가설이 기각될 경우 다른 가설도 기각될 가능성이 높아진다는 것을 보이는 것이다.

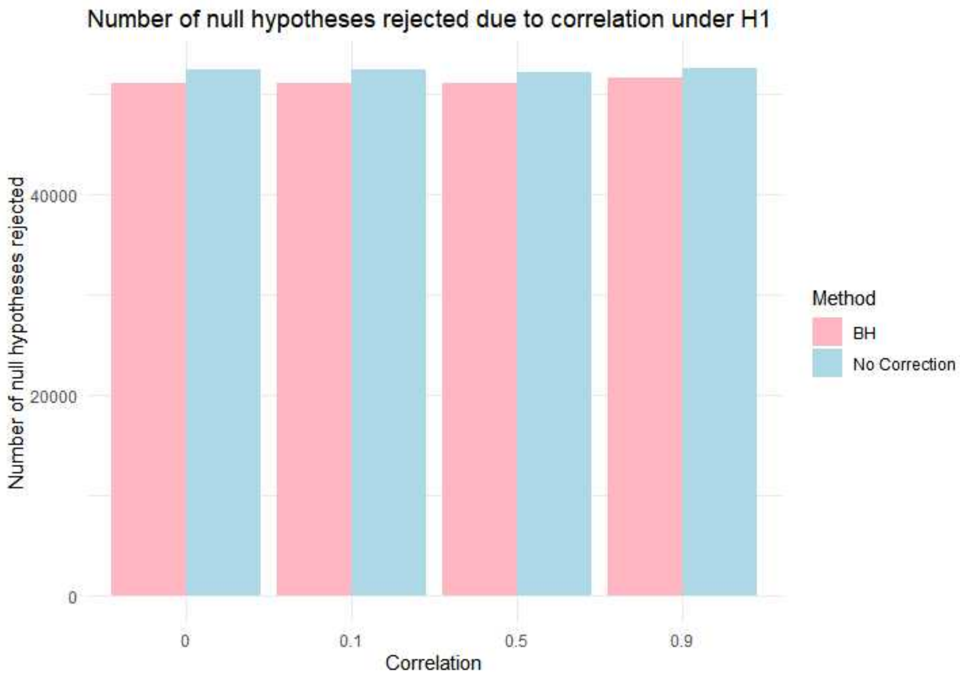


Figure5. 대립가설 하에서 상관관계에 따른 귀무가설 기각 개수

대립가설 하에서 상관관계에 따른 귀무가설 기각 개수를 나타낸 Figure5를 살펴보면, Benjamini-Hochberg procedure가 No Correction만큼 많은 귀무가설을 기각시키지만, 제1종 오류를 제어하고 있기에 조금은 더 적게 귀무가설을 기각시키고 있다는 것을 알 수 있다. 또한 검정하는 가설들 사이가 독립인 ‘Correlation = 0’의 귀무가설 기각 개수와 검정하는 가설들 사이에 양의 연관성을 가질 때의 귀무가설 기각 개수가 크게 차이가 없는걸 보아 Benjamini-Hochberg procedure가 검정하는 가설들 사이가 독립이 아니더라도 다중 검정을 보정할 수 있다는 것을 알 수 있다.

## 4. Data Analysis

다중 검정에서 효과적인 보정 방법을 찾기 위해 사용할 데이터는 자궁경부암에 관한 유전자 발현 자료이다. SUCCEED에 등록된 여성들의 128개 자궁 경부 조직 표본에서 레이저 캡처된 상피의 Affymetrix-U133-plus2.0 기반 유전자 발현 분석을 수행한 자료를 가지고 데이터 분석을 진행하였다. 해당 데이터는 24명의 정상인 사람과 28명의 자궁경부암을 가진 사람들의 유전자 발현으로 구성되어있으며 유전자의 개수는 총 20261개 이다.

이 데이터를 가지고 각 유전자 발현 별로 정상인 집단과 암 집단의 차이가 있는지 다중 검정을 시행하고자 한다. 즉 이 다중 검정의 귀무가설은 ‘정상인 집단과 암 집단의 유전자 발현이 차이가 없다’이다. 이때 각 검정들은 독립적이며, 각 자료는 정규분포를 따르지 않는 모집단에서 임의로 추출된 자료라고 가정한 후 다중 검정을 시행한다. 또한 보정 방법으로는 1. No Correction, 2. Bonferroni Method, 3. Holm’s Method, 4. Benjamini-Hochberg procedure을 이용하였다. 모든 보정 방법에서 p값이 0.05보다 작으면 귀무가설을 기

각하도록 설정하였다.

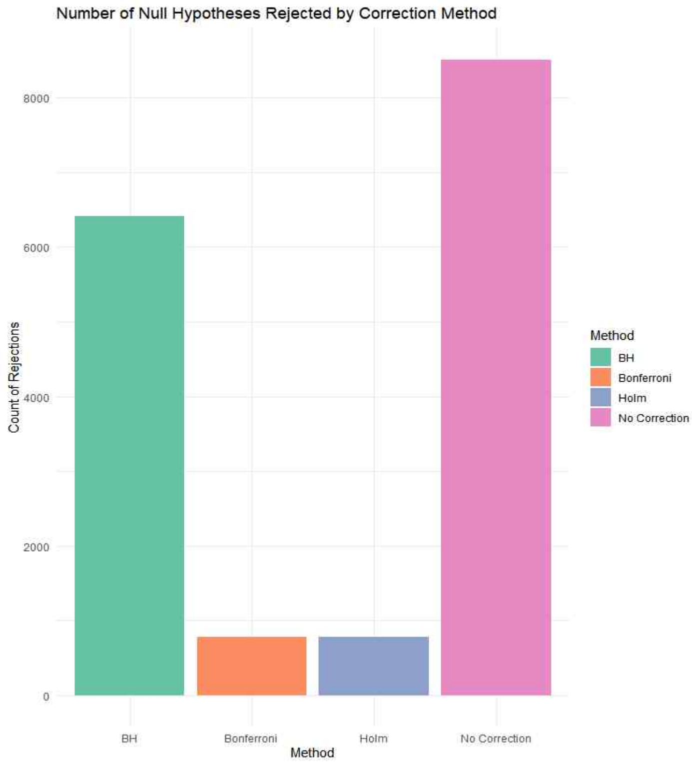


Figure6. 보정 방법별 귀무가설 기각 개수

위 다중 검정을 시행한 결과 보정 방법별로 귀무가설을 기각한 개수를 표현한 것이 바로 Figure6 이다. No Correction에서 20261개의 가설 중에서 8514개의 가설을 기각시켜 보정 방법들 중 가장 많은 귀무가설을 기각시켰다. 그 다음으로 Benjamini-Hochberg procedure에서 20261개의 가설 중에서 6417개의 가설을 기각시켰고, Bonferroni Method와 Holm's Method는 20261개의 가설 중에서 780개의 가설을 기각시켰다.

1, 2장에 따르면 다중 검정 시 보정을 하지 않으면, 귀무가설이 진실이어도 기각하는 결과를 내는 false positive가 많이 발생할 수 있다고 하였다. 그래서 보정을 통해 제1종 오류의 위험을 제어하는데 FWER을 제어하는 Bonferroni Method와 Holm's Method는 총 가설의 개수에 의존하므로 매우 보수적으로 나타난다. 즉 귀무가설을 잘 기각시키지 않아 제2종 오류로 인해 검정력이 낮아질 수 있다고 하였다. 그래서 FDR을 제어하는 방법인 Benjamini-Hochberg procedure를 사용하여 검정력이 낮아지는 것을 보완하고자 하였다. 위 Figure1을 보게 되면 FWER를 제어하는 Bonferroni Method와 Holm's Method는 매우 보수적이어서 No Correction과 Benjamini-Hochberg procedure이 기각한 귀무가설 개수보다 훨씬 더 적게 귀무가설을 기각한 것을 알 수 있다. 또한 Benjamini-Hochberg procedure는 FWER를 제어하는 방법보다는 덜 보수적이면서 제1종 오류의 위험을 제어하기에, No Correction보다는 귀무가설을 덜 기각시키지만 Bonferroni Method와 Holm's Method보다는 귀무가설을 더 기각시켰다. 결론적으로 실제 데이터에서도 다중 검정의 보정 효과가 잘 나타나는 것을 알 수 있다.

| Row | gene_index_1 | gene_index_2 | no_correction | bonferroni | holm | bh |
|---|---|---|---|---|---|---|
| 1 | 54965 | PIGX | o | o | o | o |
| 2 | 91624 | NEXN | o | o | o | o |
| 3 | 11202 | KLK8 | o | o | o | o |
| 4 | 11078 | TRIOBP | o | o | o | o |
| 5 | 121441 | NEDD1 | o | o | o | o |
| ... | ... | ... | ... | ... | ... | ... |
| 8511 | 100506652 | NA | o | x | x | x |
| 8512 | 171484 | FAM9C | o | x | x | x |
| 8513 | 126259 | TMIGD2 | o | x | x | x |
| 8514 | 100996551 | NA | o | x | x | x |
| 8515 | 166348 | KBTBD12 | o | x | x | x |

Figure7. 보정 방법별 귀무가설 기각 여부

위 Figure2는 보정 방법별 귀무가설 기각 여부에 대하여 나타낸 데이터 프레임이다. 각 보정 방법에서 수준 0.05 하에서 귀무가설을 기각시켰다면 'o' 표시를, 기각시키지 못했다면 'x' 표시를 하였다. Figure2를 통해 어떤 유전자 발현이 암환자 집단과 정상인 집단의 차이를 설명하기에 유용한지 알 수 있다.

## 5. Conclusion

본 연구에서는 다중 검정 보정 방법 접근 방식인 고전적 접근 방식과는 다르다. 고전적 접근 방식은 강한 의미에서의 FWER을 제어하는 것을 요구하는 보수적인 제1종 오류율 제어를 필요로 한다. 이에 새로운 접근 방식인 FDR 제어 방법을 제시하였다. 이 FDR 제어 방법인 Benjamini-Hochberg procedure는 덜 보수적이면서 검정력을 높여, 많은 분야에서의 테스트에서 적절하게 제1종 오류를 제어할 수 있다.

하지만 Benjamini-Hochberg procedure는 독립적인 상황에서만 사용할 수 있는데, 대부분의 연구에서는 검정 사이에 의존적 관계를 가진다. 이러한 한계를 극복하기 위해 Benjamini-Hochberg procedure에 조건을 주어 검정 사이에 양의 의존성을 가져도 다중 검정 보정을 효과적으로 할 수 있다는 결과를 얻었다.

## References

Yoav Benjamini; Yosef Hochberg."Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing" Journal of the Royal Statistical Society. Series B (Methodological), Vol.57, No.1(1995), 289-300.

Yoav Benjamini and Daniel Yekutieli. "THE CONTROL OF THE FALSE DISCOVERY RATE IN MULTIPLE TESTING UNDER DEPENDENCY" The Annals of Statistics 2001, Vol. 29, No. 4, 1165-1188.