



다양한 가설 검정 방법에서의 유의성 분석에 대한 효과크기 계산

이다은, 이은경
이화여자대학교 자연과학대학 통계학과

연구배경/연구 목적

- 연구에서 결과의 유의성을 입증하기 위해 사용되는 귀무가설(NHST)은 유의 확률(P-value)을 이용하여 그 결과를 판단한다. 그러나 유의 확률을 이용해 결과를 이분법적으로만 분류하는 것은 잘못된 해석으로 이어질 수 있다.
- 본 연구에서는 연구 결과에 대한 통계적 유의성을 정확히 해석하고 결과에 대한 자세한 내용을 제시하기 위한 방법으로 효과 크기(effect size)를 설명하는 것이다. 다양한 통계 검정 방법에 따른 효과 크기 계산 방법을 설명하고, R의 내장 데이터터를 이용하여 계산 및 결과를 비교하였다. 또한 이를 Shiny R로 구현한 어플리케이션을 제공한다.

연구방법

- t-통계검정, 회귀분석, ANOVA, 분할표 검정에 따른 적절한 효과 크기 계산 방법을 아래 세 가지 유형으로 나누어 정의하고 그 의미를 설명하였다.

r-계열

피어슨 상관계수, 스피어만 상관계수, 결정계수와 수정된 결정계수, η^2 와 partial- η^2 , Cohen's f^2 , ω^2

d-계열

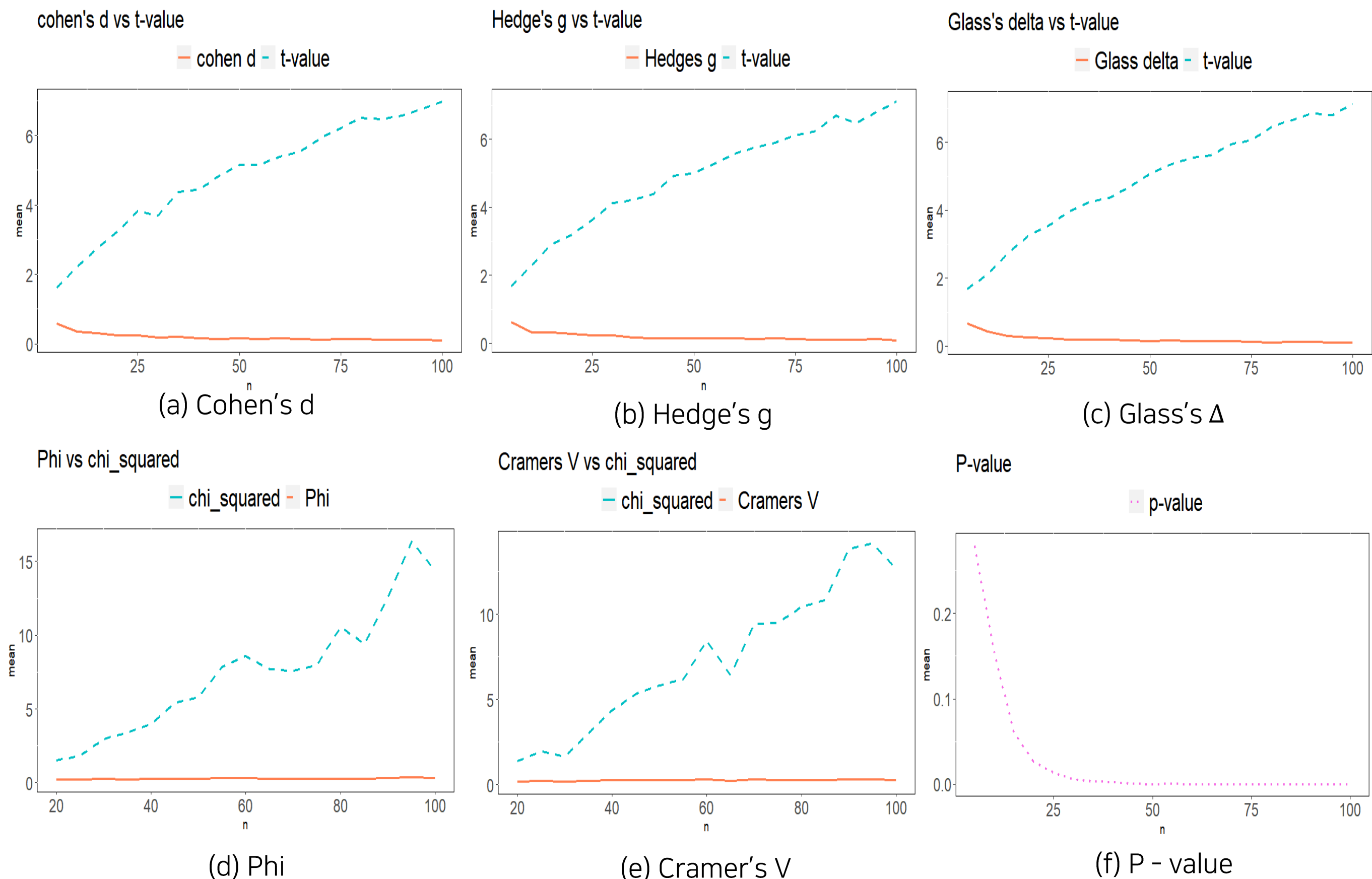
Cohen's d, Hedge's g, Glass's Δ

c-계열

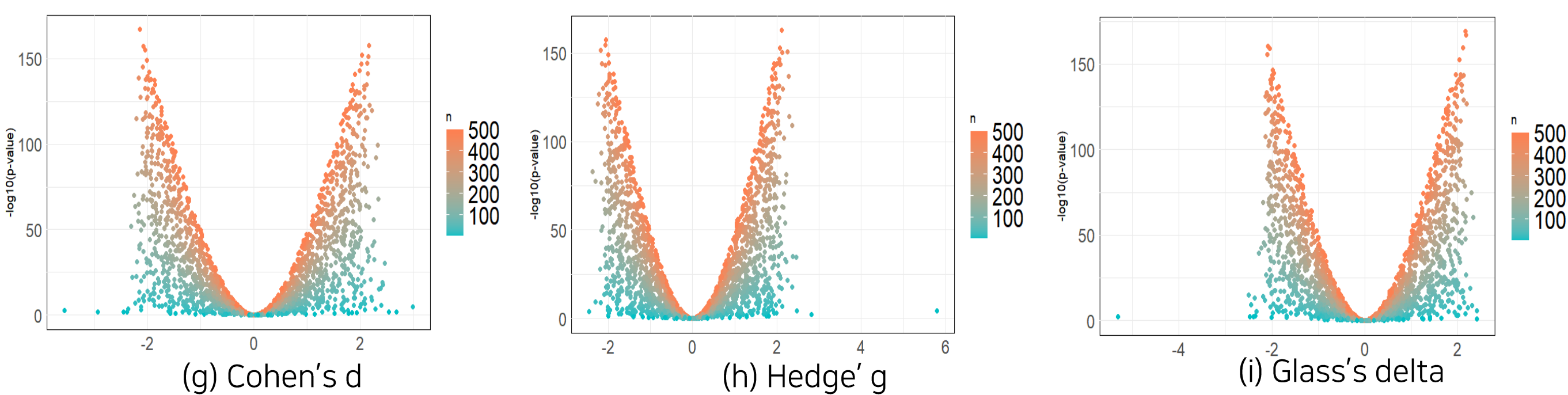
Phi, 오즈비, Cramer's V

연구결과

A. 표본 크기에 따른 각 t / F 통계량과 효과 크기, 유의확률 변화 그림



B. 표본 크기에 따른 유의확률과 효과크기의 volcano plot



I r - family

효과 크기	검정	공식	가이드라인	
Pearson' r	linear correlation	$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2} \sqrt{\sum(y - \bar{y})^2}}$	$ ES < 0.1$	very Small
			$0.1 \leq ES < 0.3$	Small
			$0.3 \leq ES < 0.5$	Medium
Spearman's ρ	rank correlation	$\rho = \frac{6 \sum(u-v)^2}{N(N^2-1)}$	$ ES \geq 0.5$	Large
R^2	simple linear regression	$R^2 = \frac{(\sum(x - \bar{x})(y - \bar{y}))^2}{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}$	$R^2 < 0.02$	very Small
			$0.02 \leq R^2 < 0.13$	Small
			$0.13 \leq R^2 < 0.26$	Medium
Adj - R^2	multiple linear regression	$R^2_{adj} = R^2 - \frac{(1 - R^2)p}{N - p - 1}$	$R^2 \geq 0.26$	Large
η^2	1-way ANOVA	$\eta^2 = \frac{SS_{factor}}{SS_{total}}$	$ES < 0.01$	very Small
			$0.01 \leq ES < 0.06$	Small
			$0.06 \leq ES < 0.14$	Medium
Partial - η^2	n-way ANOVA	$partial \eta^2 = \frac{SS_{factor}}{SS_{factor} + SS_{error}}$	$ES \geq 0.14$	Large
Cohen's f^2	multiple linear regression / n-way ANOVA	$f^2 = \frac{R^2}{1 - R^2}$ $f^2 = \frac{\eta^2}{1 - \eta^2}$	$f^2 < 0.02$	very Small
			$0.02 \leq f^2 < 0.15$	Small
			$0.15 \leq f^2 < 0.35$	Medium
ω^2	1-way / n-way ANOVA	$\omega^2 = \frac{SS_{factor} - (k - 1)MSE}{SS_{total} + MSE}$	$\omega^2 < 0.02$	very Smal
			$0.02 \leq \omega^2 < 0.13$	Small
			$0.13 \leq \omega^2 < 0.26$	Medium
			$\omega^2 \geq 0.26$	Large

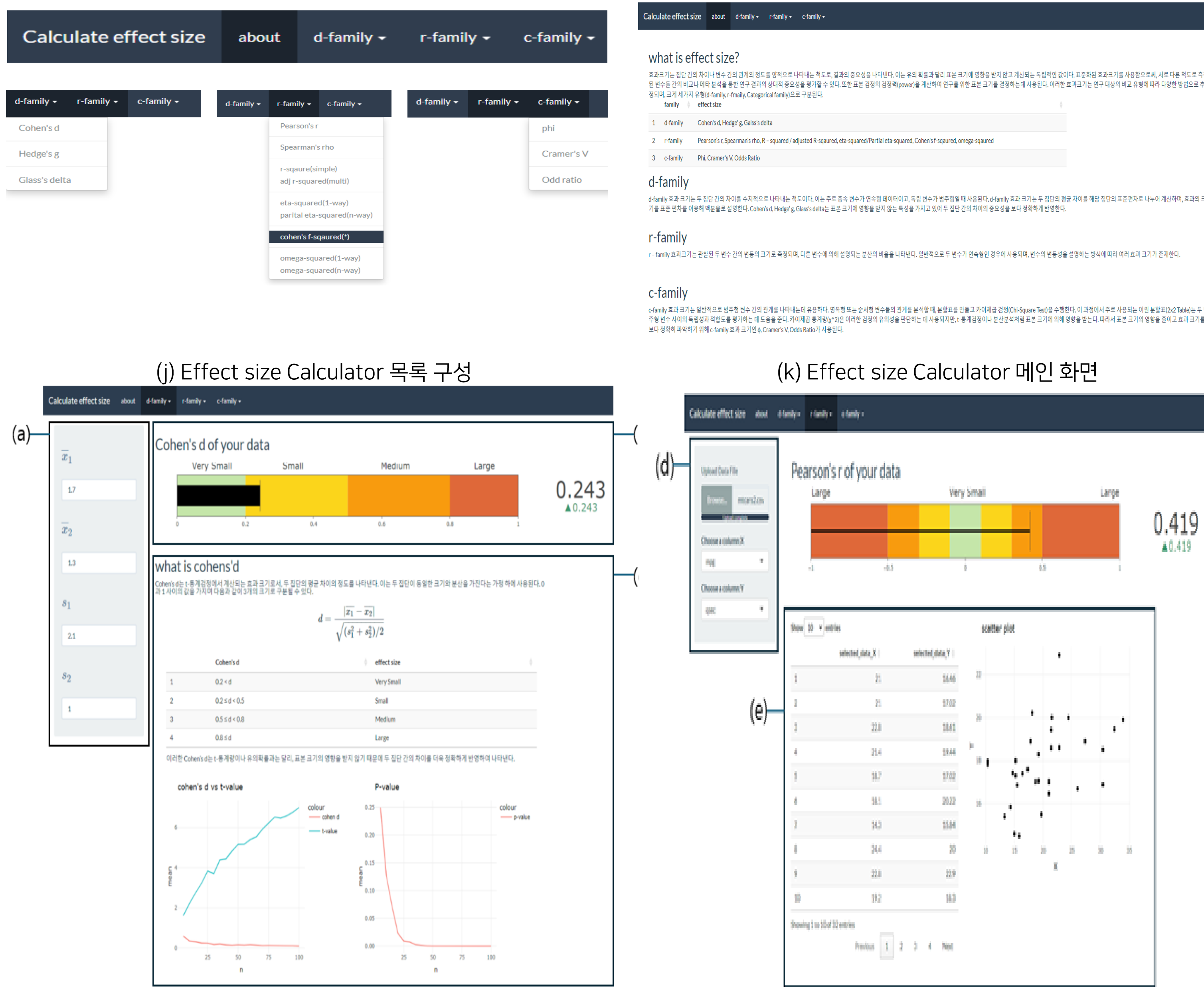
II d - family

효과 크기	검정	공식	가이드라인	
Cohen's d	t-test With equal sample size / variance	$d = \frac{\bar{X}_1 - \bar{X}_2}{S_{within}}$ $S_{within} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$	$ ES < 0.2$	very Small
			$0.2 \leq ES < 0.5$	Small
Hedge's g	t-test With unequal / small sample size	$g = \left(1 - \frac{3}{4df-1}\right) \times d$	$0.5 \leq ES < 0.8$	Medium
			$ ES \geq 0.8$	Large
Glass's Δ	t-test With unequal Variance	$\Delta = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_{control}^2}}$		

III c - family

효과 크기	검정	공식	가이드라인	
Phi	2X2 table	$\phi = \sqrt{\frac{\chi^2}{N}}$	$ES < 0.1$	very Small
			$0.1 \leq ES < 0.3$	Small
Cramer's V	mXn table	$V = \sqrt{\frac{\chi^2}{N(m-1)}}$	$0.3 \leq ES < 0.5$	Medium
			$ES \geq 0.5$	Large
Odds ratio	2X2 table	$OR = \frac{(x_1y_1)(x_0y_0)}{(x_1y_0)(x_0y_1)} = \frac{ad}{bc}$	$OR < 1.44$	very Small
			$1.44 \leq OR < 2.48$	Small
			$2.48 \leq OR < 4.27$	Medium
			$OR \geq 4.27$	Large

IV Effect size Calculator (<https://shinnylee.shinyapps.io/EffectSizeCalculator/>)



(l) Effect size Calculator 각 효과크기 화면 구

* (a) 통계량 입력 상자, (b) 범위 막대, (c) 설명 화면(효과크기 이론적 설명)

* (d) 통계량 입력 상자(데이터 업로드 방식) (e) 설명 화면(데이터 테이블 및 산점도)

결론 및 제언

- 효과 크기는 연구 결과의 유의성의 정도를 나타내며, 유의수준과 함께 제시되어 결과의 실질적인 의미를 해석하는데 도움이 된다. 또한 표준화된 효과 크기는 다양한 연구 간의 결과를 비교할 수 있게 하며, 연구에 필요한 적절한 표본 크기를 결정하는 데에도 사용된다.
- 그러나 해석은 가이드라인에 맞춰 엄격하게 진행되는 것이 아닌 연구의 맥락, 사용된 데이터, 분석의 목적 등을 고려하여 유연하게 해석되어야 한다. 결과의 보다 자세한 설명을 위해서는 신뢰구간을 함께 제시하는 것이 권장된다.

참고문헌

- Glaser, B. E. (2023). ESDist (v1.0.0) <https://doi.org/10.5281/zenodo.10040656>.
- Cohen, J. (1988). "Statistical Power Analysis for the Behavioral Sciences" (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Pearson, K. (1896). "Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia". Philosophical Transactions of the Royal Society of London, Series A, 187, 253-318.
- Olejnik, S., & Algina, J. (2003). "Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs". Psychological Methods, 8(4), 434-447.
- Spearman, C. (1904). "The proof and measurement of association between two things". The American Journal of Psychology, 15(1), 72-101.