

이 문서는 Minitab 통계 소프트웨어의 보조 도구에서 사용되는 방법과 데이터 검사를 개발하기 위해 Minitab 통계 학자들이 실시한 연구에 대해 설명하는 전체 백서 중 하나입니다.

1-표본 t-검정

개요

1-표본 t-검정은 공정 평균을 추정하고 이 평균을 목표값과 비교하기 위해 사용됩니다. 이 검정은 표본이 어느 정도 클 때 정규성 가정의 영향을 거의 받지 않기 때문에 로버스트한 절차로 간주됩니다. 대부분의 통계 교과서에 따르면 1-표본 t-검정과 t-평균 신뢰 구간은 크기가 30개 이상인 표본에 적합합니다.

본 문서에서는 최소 30개 표본 단위에 대한 이 일반적인 규칙을 평가하기 위해 실시한 시뮬레이션에 대해 설명합니다. 해당 시뮬레이션에서는 정규성이 1-표본 t-검정에 미치는 영향에 초점을 맞췄습니다. 비정상 데이터가 검정 결과에 미치는 영향도 평가하고자 했습니다.

당사에서 실시한 연구를 토대로, 보조 도구는 다음 검사를 데이터에 대해 자동으로 수행하고 보고서 카드에 결과를 표시합니다.

- 비정상 데이터
- 정규성(정규성이 문제가 되지 않을 만큼 표본 크기가 큰가?)
- 표본 크기

1-표본 t-검정에 대한 전체적인 내용은 Arnold(1990), Casella & Berger(1990), Moore & McCabe(1993) 및 Srivastava(1958)를 참조하십시오.

참고 본 문서의 결과는 보조 도구의 쌍체 t-검정에도 적용됩니다. 쌍체 t-검정은 1-표본 t-검정 방법을 쌍체 차이 표본에 적용하기 때문입니다.

데이터 검사

비정상 데이터

비정상 데이터는 특이치라고도 하는 극도로 크거나 작은 데이터 값입니다. 비정상 데이터는 분석 결과에 중대한 영향을 미칠 수 있습니다. 표본이 작을 때 비정상 데이터는 통계적으로 유의한 결과를 찾을 확률에 영향을 미칠 수 있습니다. 비정상 데이터는 데이터 수집에 문제가 있음을 나타내거나, 연구하는 프로세스의 비정상적인 동작에 기인할 수 있습니다. 이런 데이터 점은 조사할만한 가치가 많으며, 가능하면 수정해야 합니다.

목적

Minitab에서는 전체 표본에 비해 매우 크거나 매우 작고 분석 결과에 영향을 미칠 수 있는 데이터 값이 있는지 확인하는 방법을 개발하고자 했습니다.



방법

Minitab에서는 Hoaglin, Iglewicz & Tukey(1986)가 설명한 상자 그림에서 특이치를 식별하는 방법을 토대로 비정상 데이터를 검사하는 방법을 개발했습니다.

결과

보조 도구는 분포의 하위 또는 상위 사분위수를 벗어난 사분위간 범위의 1.5배 이상인 데이터 점을 비정상 데이터로 구분합니다. 하위 및 상위 사분위수는 데이터의 25번째 및 75번째 백분위수에 해당됩니다. 사분위간 범위는 두 사분위수의 차이입니다. 이 방법을 사용하면 특정 특이치를 각각 탐지할 수 있기 때문에 특이치가 여러 개일 때도 효과적입니다.

비정상 데이터가 있는지 검사할 때 1-표본 t-검정에 대해 보조 도구 보고서 카드에는 다음과 같은 상태가 표시됩니다.

상태	조건
	비정상적인 데이터 점 없음.
	하나 이상의 데이터 점이 비정상이고 검정 결과에 영향을 미칠 수 있음.

정규성

1-표본 t-검정은 모집단이 정규 분포를 따른다는 가정 하에 도출됩니다. 데이터가 정규 분포를 따르지 않더라도 표본이 충분히 클 경우에는 이 방법이 효과적입니다.

목적

Minitab에서는 비정규성이 검정의 제1종 오류와 제2종 오류에 미치는 영향을 확인하여 표본 크기와 정규성에 대한 기준치를 제시하고자 했습니다.

방법

1-표본 t-검정을 수행하거나 모집단의 평균에 대해 t-신뢰 구간을 계산할 때 정규성 가정을 무시할 수 있는 표본 크기를 확인하기 위한 시뮬레이션을 실시했습니다.



정규성이 검정의 제1종 오류율에 미치는 영향을 평가하기 위한 첫 번째 실험을 설계했습니다. 구체적으로, 검정이 모집단 분포의 영향을 받지 않기 위해 필요한 최소 표본 크기를 유추하고자 했습니다. 정규 및 비정규 모집단에서 생성한 소규모, 중간 및 대규모 표본에 대해 1-표본 t-검정을 실시했습니다. 비정규 모집단에는 조금 또는 많이 치우친 모집단과 꼬리가 두껍고 얇은 대칭 모집단, 그리고 오염된 정규 모집단이 포함되었습니다. 정규 모집단은 비교를 위한 통제 모집단으로 사용되었습니다. 각 경우에 대해 모의 유의 수준을 계산하고 목표 또는 명목 유의 수준 0.05와 비교했습니다. 검정이 효과적이면 모의 유의 수준이 0.05와 가까울 것입니다. 다양한 조건에서 모두 모의 유의 수준을 조사하여 분포에 관계 없이 목표 수준과 가깝게 유지되는 최소 표본 크기를 평가했습니다. 자세한 내용은 부록 A를 참조하십시오.

두 번째 실험에서는 정규성이 검정의 제2종 오류율에 미치는 영향을 조사했습니다. 시뮬레이션의 설계는 첫 연구와 동일합니다. 그러나 다양한 조건에서 관측된 모의 검정력 수준을 1-표본 t-검정의 이론적 검정력 함수를 사용하여 계산한 목표 검정력 수준에 비교했습니다. 자세한 내용은 부록 B를 참조하십시오.

결과

비정규성이 검정의 제1종 및 제2종 오류율에 모두 미치는 영향은 표본 크기가 20개 이하일 경우 미미했습니다. 그러나 표본의 모집단이 극도로 치우친 분포를 따를 경우에는 더 큰 표본이 필요할 수 있습니다. 이런 경우에는 약 40개의 표본 크기가 권장됩니다. 자세한 내용은 부록 A와 부록 B를 참조하십시오.

이 검정은 표본이 상대적으로 작을 때도 효과적이기 때문에 보조 도구는 데이터의 정규성을 검사하지 않습니다. 대신 표본 크기를 확인하고 보고서 카드에 다음 상태를 표시합니다.

상태	조건
	표본 크기가 20개 이상이므로 정규성은 문제되지 않음.
	표본 크기가 20개 미만이므로 정규성이 문제될 수 있음.

표본 크기

일반적으로, 가설 검정은 "차이가 없는" 귀무 가설을 기각하기 위한 증거를 수집하기 위해 실시합니다. 표본이 너무 작으면 검정의 검정력이 실제로 존재하는 평균의 차이를 탐지하는데 충분하지 않아서 제2종 오류가 발생할 수 있습니다. 따라서 실제로 중요한 차이를 높은 확률로 탐지할 수 있을 만큼 큰 표본 크기를 사용하는 것이 중요합니다.

목적

데이터에서 귀무 가설을 반박할 증거를 충분히 얻을 수 없는 경우, 표본 크기가 검정에서 탐지하려는 실제 차이를 높은 확률로 탐지하기에 충분히 크지 확인하고자 했습니다. 표본 크기 계획의 목적은 중요한 차이를 높은 확률로 탐지하기에 충분히 큰 표본 크기를 사용하도록 보장하는 것이지만, 표본이 너무 커서 무의미한 차이가 통계적으로 중요하다고 인식될 확률이 높아져서도 안 됩니다.

방법

검정력 및 표본 크기 분석은 통계적 분석을 실시하는 데 사용되는 특정 검정의 이론적 검정력 함수에 기초합니다. 앞에서 설명한 대로 1-표본 t-검정의 검정력 함수는 표본 크기가 20개 이상일 때 정규 가정의 영향을 받지 않습니다. 검정력 함수는 표본 크기와 목표 평균값의 차이, 그리고 모집단의 평균과 모집단의 분산에 의해 결정됩니다. 자세한 내용은 부록 B를 참조하십시오.






결과

데이터로부터 귀무 가설을 반박하는 충분한 증거를 얻을 수 없으면 보조 도구는 주어진 표본 크기에 대해 80% 및 90% 확률로 탐지할 수 있는 실제 차이를 계산합니다. 또한 탐지하려는

특정 실제 차이를 사용자가 제공할 경우 보조 도구는 차이를 80% 및 90% 확률로 탐지할 수 있는 표본 크기를 계산합니다.

결과는 구체적인 표본에 따라 달라지므로 보고할 일반적인 결과는 없습니다. 그러나 1-표본 t-검정의 검정력에 대한 자세한 내용은 부록 B를 참조하십시오.

검정력과 표본 크기를 확인할 때 1-표본 t-검정에 대한 보조 도구 보고서 카드에는 다음과 같은 상태가 표시됩니다.

상태	조건
	검정에서 평균과 목표값의 차이를 찾으므로 검정력은 문제가 되지 않음. 또는 검정력이 충분함. 검정에서 평균과 목표값의 차이를 찾지 못했지만 표본 크기는 주어진 차이를 90% 이상의 확률로 탐지하기에 충분함.
	검정력이 충분할 수 있음. 검정에서 평균과 목표값의 차이를 찾지 못했지만 표본 크기는 주어진 차이를 80%에서 90%의 확률로 탐지하기에 충분함. 90% 검정력을 달성하기 위해 필요한 표본 크기가 보고됨.
	검정력이 충분하지 않을 수 있음. 검정에서 평균과 목표값의 차이를 찾지 못했으며 표본 크기는 주어진 차이를 60%에서 80%의 확률로 탐지하기에 충분함. 80% 검정력과 90% 검정력을 달성하는 데 필요한 표본 크기가 보고됨.
	검정력이 충분하지 않음. 검정에서 평균과 목표값의 차이를 찾지 못했으며 표본 크기는 주어진 차이를 60% 이상의 확률로 탐지하기에 충분하지 않음. 80% 검정력과 90% 검정력을 달성하는 데 필요한 표본 크기가 보고됨.
	검정에서 평균과 목표값의 차이를 찾지 못함. 탐지할 평균과 목표값의 실제 차이를 지정하지 않았으므로 보고서에는 표본 크기, 표준 편차 및 알파에 따라 80% 및 90% 확률로 탐지할 수 있는 차이가 표시됨.

참고 문헌

Arnold, S.F. (1990). *Mathematical statistics*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Casella, G., & Berger, R. L. (1990). *Statistical inference*. Pacific Grove, CA: Wadsworth, Inc.

Hoaglin, D. C., Iglewicz, B. & Tukey, J. W. (1986). Performance of some resistant rules for outlier labeling. *Journal of the American Statistical Association*, 81, 991-999.

Moore, D.S. & McCabe, G.P. (1993). *Introduction to the practice of statistics*, 2nd ed. New York, NY: W. H. Freeman and Company.

Neyman, J., Iwazskiewicz, K. & Kolodziejczyk, S. (1935). Statistical problems in agricultural experimentation, *Journal of the Royal Statistical Society, Series B*, 2, 107-180.

Pearson, E.S., & Hartley, H.O. (Eds.). (1954). *Biometrika tables for statisticians*, Vol. I. London: Cambridge University Press.

Srivastava, A. B. L. (1958). Effect of non-normality on the power function of t-test, *Biometrika*, 45, 421-429.

부록 A: 비정규성이 유의 수준(검정의 유효성)에 미치는 영향

정규 가정 하에서 1-표본 t-검정은 UMP(uniformly most powerful) 불편 α 크기 검정입니다. 즉, 이 검정은 평균에 대한 다른 모든 불편 α 크기 검정 이상으로 강력합니다. 그러나 표본의 모집단이 정규 분포를 따르지 않으면 위의 최적 속성은 표본 크기가 충분히 클 경우에만 참입니다. 다시 말해 표본이 충분히 클 경우 1-표본 t-검정의 실제 유의 수준은 정규 및 비정규 데이터에 대한 목표 수준과 대략적으로 같고 검정의 검정력 함수도 정규 가정의 영향을 받지 않습니다(Srivastava, 1958).

당사는 t-검정이 정규 가정의 영향을 받지 않을 만큼 표본이 크다고 간주되려면 표본이 얼마나 커야 하는지 확인하고자 했습니다. 많은 교과서에서는 표본 크기가 $n \geq 30$ 일 경우 대부분의 실제 용도에서 정규 가정을 무시할 수 있다고 권장합니다(Arnold, 1990; Casella & Berger, 1990; Moore & McCabe, 1993). 본 문서의 부록에서 설명하는 조사의 목적은 다양한 비정규 분포가 1-표본 t-검정에 미치는 영향을 조사하여 이 일반 규칙을 평가하기 위한 시뮬레이션 연구를 실시하는 것입니다.

시뮬레이션 연구 A

당사는 비정규성이 검정의 제1종 오류율에 미치는 영향을 조사하여 오류율이 분포에 관계 없이 안정적이고 목표 오류율과 가깝게 유지되는 최소 표본 크기를 산정하고자 했습니다.

이를 위해 $\alpha = 0.05$ 를 사용한 양측 t-검정을 속성이 서로 다른 여러 분포로부터 생성된 다양한 크기($n = 10, 15, 20, 25, 30, 35, 40, 50, 60, 80, 100$)의 무작위 표본을 사용하여 수행했습니다. 사용된 분포는 다음과 같습니다.

- 표준 정규 분포($N(0,1)$)
- 자유도가 5-10인 t-분포($t(5), t(10)$)처럼 꼬리가 두꺼운 대칭 분포
- 위치가 0이고 척도가 1인 Laplace 분포(Lpl)
- 척도가 1인 지수 분포(Exp)와 자유도가 3, 5, 10인 카이-제곱 분포($Chi(3), Chi(5), Chi(10)$)로 대표되는 치우친 두꺼운 꼬리 분포

- 균등 분포(U(0,1))와 두 모수가 모두 3으로 설정된 베타 분포(B(3,3))처럼 꼬리가 얇은 대칭 분포
- 왼쪽으로 치우친 두꺼운 꼬리 분포(B(8,1))

또한 특이치의 직접적인 영향을 평가하기 위해 다음과 같이 정의되는 오염된 정규 분포로부터 표본을 생성했습니다.

$$CN(p, \sigma) = pN(0,1) + (1 - p)N(0, \sigma)$$

여기서 p 는 혼합 모수로 정의되고 $1 - p$ 는 오염의 비율, 즉 특이치의 비율입니다. 연구에 사용할 오염된 정규 분포 2개를 선택했습니다. 하나는 모집단 구성원의 10%가 특이치인 $CN(0.9,3)$ 이었고 다른 하나는 모집단 구성원의 20%가 특이치인 $CN(.8,3)$ 이었습니다. 이 두 분포는 대칭이며, 특이치로 인해 꼬리가 길니다.

표본 크기별로 10,000개의 표본 반복실험을 각 모집단에서 추출하고 귀무 가설 $\mu = \mu_0$ 와 대립 가설 $\mu \neq \mu_0$ 를 사용한 1-표본 t-검정을 표본 10,000개에 대해 각각 수행했습니다. 각 검정에 대해 가설 평균 μ_0 를 표본 모집단의 실제 평균으로 설정했습니다. 그 결과, 주어진 표본 크기에 대해 10,000번의 표본 반복실험 중 귀무 가설이 기각되는 결과를 가져오는 횟수의 비율에 해당되는 부분이 검정의 모의 제1종 오류율 또는 유의 수준을 나타냅니다. 목표 유의 수준이 5%이므로, 시뮬레이션 오류는 약 0.2%입니다.

시뮬레이션 결과는 표 1과 2에 나와 있으며, 그림 1과 2에 그래프로 표시되어 있습니다.

결과 및 요약

결과(표 1 및 그림 1 참조)는 표본을 대칭 모집단으로부터 생성할 경우 표본 크기가 작을 때도 검정의 모의 유의 수준이 목표 유의 수준에 가까움을 보여줍니다. 그러나 표본(오염된 정규 분포에서 추출된 작은 표본 포함)이 작으면 두꺼운 꼬리 대칭 분포에 대한 검정 결과가 약간 보수적입니다. 또한 표본이 작을 때 특이치가 있으면 검정의 유의 수준이 감소하는 것으로 보입니다. 그러나 꼬리가 더 얇은 대칭 모집단(베타 (3,3) 및 균등 분포)에서 작은 표본을 추출한 경우에는 반대의 효과가 나타납니다. 이런 경우에는 모의 유의 수준이 약간 더 높습니다.

표 1 대칭 모집단에서 추출한 표본에 대한 양측 1-표본 t-검정의 모의 유의 수준 목표 유의 수준은 입니다.

분포	N(0,1)	t(5)	t(10)	Lpl	CN(.9,3)	CN(.8,3)	B(3,3)	U(0,1)
N	대칭 및 두꺼운 꼬리						대칭 및 얇은 꼬리	
10	0.050	0.046	0.048	0.044	0.043	0.039	0.057	0.057
15	0.051	0.050	0.049	0.049	0.043	0.043	0.053	0.054
20	0.047	0.051	0.051	0.047	0.043	0.044	0.051	0.052
25	0.050	0.047	0.050	0.046	0.046	0.046	0.048	0.050
30	0.053	0.050	0.048	0.043	0.049	0.046	0.050	0.048
35	0.052	0.047	0.049	0.050	0.047	0.045	0.051	0.054
40	0.046	0.052	0.054	0.048	0.046	0.049	0.044	0.050
50	0.050	0.049	0.051	0.048	0.047	0.051	0.053	0.050
60	0.052	0.049	0.053	0.050	0.051	0.056	0.054	0.052
80	0.049	0.050	0.051	0.047	0.047	0.052	0.049	0.049
100	0.050	0.052	0.049	0.051	0.052	0.054	0.051	0.054

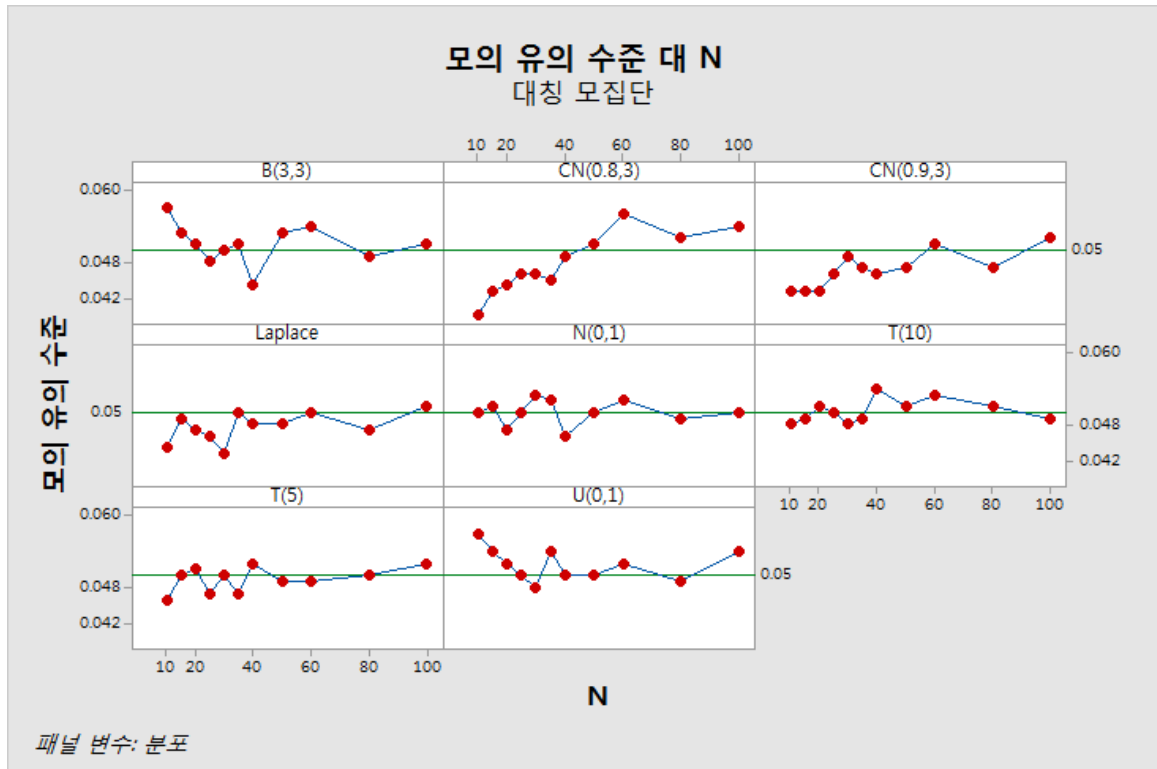


그림 1 대칭 모집단에서 생성한 표본 크기와 비교한 양측 1-표본 t-검정의 모의 유의 수준
그림 목표 유의 수준은 0.05입니다.

반면에, 표본을 치우친 분포에서 생성하면 검정의 성능이 치우친 정도에 따라 결정됩니다. 표 2와 그림 2의 결과에는 표본이 작을 때 1-표본 t-검정이 치우침의 영향을 받는 것으로 나타났습니다. 심하게 치우친 모집단(지수, Chi(3) 및 베타(8,1))의 경우 모의 유의 수준이 목표 유의 수준과 가까우려면 더 큰 표본이 필요합니다. 그러나 약간 치우친 모집단(Chi(5) 및 Chi(10))의 경우 모의 유의 수준이 목표 수준과 가까우려면 최소 표본 크기가 20이면 충분합니다. 표본 크기가 20개일 경우 자유도가 5인 카이-제곱 분포에 대한 모의 유의 수준은 약 0.063이었고 자유도가 10인 카이-제곱 분포에 대해서는 약 0.056이었습니다.

표 2 치우친 모집단에서 추출한 표본에 대한 양측 1-표본 t-검정의 모의 유의 수준. 목표 유의 수준은 입니다.

N	지수	Chi(3)	B(8,1)	Chi(5)	Chi(10)
	모집단 왜도				
	2.0	1.633	-1.423	1.265	0.894
	모의 유의 수준				
10	0.101	0.089	0.087	0.069	0.060
15	0.088	0.076	0.072	0.068	0.057
20	0.083	0.073	0.069	0.063	0.056
25	0.075	0.068	0.067	0.067	0.056
30	0.069	0.067	0.066	0.058	0.054
35	0.075	0.067	0.062	0.062	0.056
40	0.067	0.067	0.061	0.059	0.056
50	0.064	0.057	0.062	0.057	0.054
60	0.063	0.056	0.061	0.054	0.055
80	0.059	0.058	0.053	0.052	0.052
100	0.060	0.055	0.055	0.047	0.053

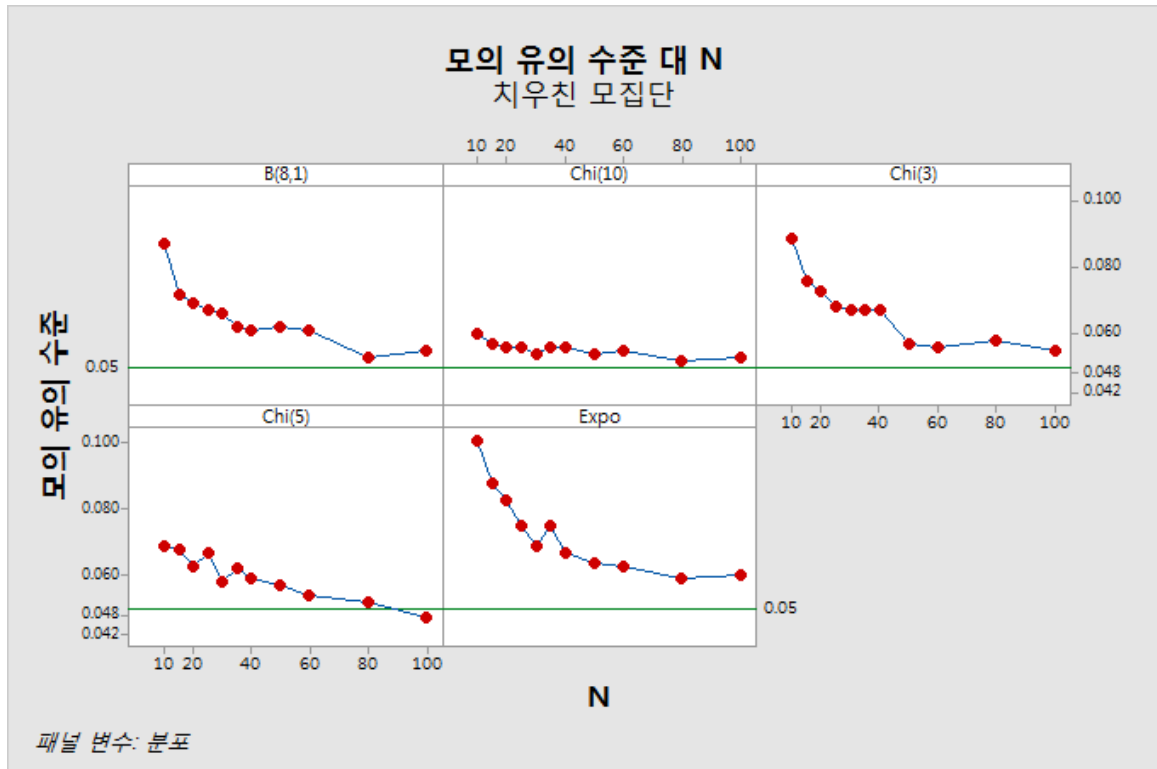


그림 2 치우친 모집단에서 추출한 표본 크기와 비교한 양측 1-표본 t-검정의 모의 유의 수준 그림. 목표 유의 수준은 0.05입니다.

이 조사에서는 신뢰 구간 대신 가설 검정에 초점을 맞췄습니다. 그러나 가설 검정과 신뢰 구간은 모두 통계적 유의성을 결정하는 데 사용할 수 있기 때문에 결과는 신뢰 구간까지 확대 적용됩니다.

부록 B: 표본 크기와

검정의 검정력

당사는 검정력 함수가 해당 함수를 도출하기 위해 사용된 정규 가정의 영향을 얼마나 받는지 조사하고자 했습니다. β 가 검정의 제2종 오류면 $1 - \beta$ 는 검정의 검정력입니다. 따라서 계획된 표본 크기는 제2종 오류율이 작거나 검정력 수준이 합당하게 높도록 결정됩니다.

t-검정의 검정력 함수는 잘 알려져 있고 잘 문서화되어 있습니다. Pearson & Hartley(1954)와 Neyman, Iwazkiewicz & Kolodziejczyk(1935)은 검정력 함수의 차트 및 표를 제공합니다.

α 크기의 양측 1-표본 t-검정의 경우 이 함수의 수학적식은 표본 크기와 실제 평균 μ 와 가설 평균 μ_0 의 차이 δ 를 사용하여 다음과 같이 표현할 수 있습니다.

$$\pi(n, \delta) = 1 - F_{n-1, \lambda}(t_{n-1}^{\alpha/2}) + F_{n-1, \lambda}(-t_{n-1}^{\alpha/2})$$

여기서 $F_{d, \lambda}(\cdot)$ 는 자유도가 $d = n - 1$ 이고 비중심 모수가 다음과 같은 비중심 t 분포의 C.D.F입니다.

$$\lambda = \frac{\delta\sqrt{n}}{\sigma}$$

그리고 t_d^α 는 자유도가 d 인 t 분포의 100α 상위 백분위 점을 나타냅니다.

단측 대립 가설의 경우, 검정력의 식은

$$\pi(n, \delta) = 1 - F_{n-1, \lambda}(t_{n-1}^\alpha)$$

(귀무 가설을 $\mu > \mu_0$ 에 대해 검정할 때), 그리고

$$\pi(n, \delta) = F_{n-1, \lambda}(-t_{n-1}^\alpha)$$

(귀무 가설을 $\mu < \mu_0$ 에 대해 검정할 때)로 표현됩니다.

이러한 함수는 데이터가 정규 분포를 따르고 검정의 유의 수준이 α 값으로 고정되었다는 가정 하에 도출됩니다.

시뮬레이션 연구 B

이 시뮬레이션은 비정규성이 1-표본 t-검정의 이론적 검정력 함수에 미치는 영향을 평가하도록 설계되었습니다. 정규성의 영향을 평가하기 위해 검정의 이론적 검정력 함수를 사용하여 계산한 목표 검정력 수준과 모의 검정력 수준을 비교했습니다.

당사는 첫 번째 시뮬레이션 연구(부록 A 참조)에서 설명한 것과 동일한 모집단에서 추출한 다양한 크기($n = 10, 15, 20, 25, 30, 35, 40, 50, 60, 80, 100$)의 무작위 표본에 대해 $\alpha = 0.05$ 에서 양측 t-검정을 수행했습니다.

각 모집단에 대해 검정의 귀무 가설은 $\mu = \mu_0 - \delta$ 이고 그 대립 가설은 $\neq \mu_0 - \delta$ 입니다. 여기서 μ_0 는 모집단의 실제 평균으로 설정되고 $\delta = \sigma/2$ (σ 는 모집단의 표준 편차)입니다. 따라서 실제 평균과 가설 평균의 차이는 0이므로 귀무 가설을 기각하는 것이 올바른 결정입니다.

주어진 각 표본 크기에 대해 각 분포에서 10,000개의 표본 반복실험을 추출했습니다. 주어진 각 표본 크기에 대해 10,000회의 반복실험 중에서 귀무 가설이 기각된 비율의 부분이 주어진 표본 크기와 차이 δ 에 대한 검정의 모의 검정력 수준을 나타냅니다. 이 차이값을 선택한 이유는 표본 크기가 작을 때 상대적으로 작은 검정력 값을 생성하기 때문입니다.

또한 목표 검정력 값이라고 하는 해당 이론적 검정력 값을 모의 검정력 값과 비교하기 위해 차이 δ 와 다양한 표본 크기에서 계산합니다.

시뮬레이션 결과는 표 3과 4에 나와 있으며, 그림 3과 4에 그래프로 표시되어 있습니다.

결과 및 요약

결과에서는 표본 크기가 충분히 크면 1-표본 t-검정의 검정력이 대체로 정규 가정의 영향을 받지 않는 것으로 확인되었습니다. 대칭 모집단에서 생성한 표본에 대해 표 3에는 표본 크기가 작은 경우에도 목표 검정력과 모의 검정력 수준이 가깝다는 결과가 나와 있습니다. 그림 3에 표시된 해당 검정력 곡선은 실제로 구분할 수 없습니다. 오염된 정규 분포에서 생성한 표본에 대해 검정력 값은 표본 크기가 중간 이하일 경우 다소 보수적입니다. 그 이유는 해당 모집단에 대한 검정의 실제 유의 수준이 고정된 목표 유의 수준 보다 약간 더 높기 때문일 수 있습니다.

.표 3 표본을 대칭 모집단에서 추출했을 때 차이 $\delta = \sigma/2$ 에서 크기가 $\alpha = 0.05$ 인 양측 1-표본 t-검정의 모의 검정력 수준 모의 검정력 수준을 정규 가정 하에서 도출된 이론적 목표 검정력 수준과 비교합니다.

n	목표 검정력	N(0,1)	t(5)	t(10)	Lpl	CN(.9,3)	CN(.8,3)	B(3,3)	U(0,1)
		$\delta = \sigma/2$ (대칭 모집단)일 때 모의 검정력 수준							
10	0.293	0.299	0.334	0.311	0.357	0.361	0.385	0.28	0.269
15	0.438	0.438	0.48	0.45	0.491	0.512	0.511	0.423	0.421
20	0.565	0.57	0.603	0.578	0.60	0.629	0.623	0.557	0.548
25	0.67	0.674	0.695	0.68	0.691	0.712	0.70	0.665	0.67
30	0.754	0.756	0.77	0.756	0.767	0.768	0.765	0.754	0.75
35	0.82	0.819	0.827	0.815	0.82	0.819	0.812	0.822	0.818
40	0.869	0.87	0.871	0.868	0.862	0.869	0.868	0.875	0.867
50	0.934	0.933	0.929	0.93	0.929	0.923	0.925	0.932	0.94
60	0.968	0.967	0.963	0.965	0.964	0.955	0.955	0.968	0.971
80	0.993	0.993	0.989	0.992	0.991	0.988	0.989	0.994	0.994
100	0.999	0.998	0.996	0.998	0.999	0.998	0.996	0.999	0.999

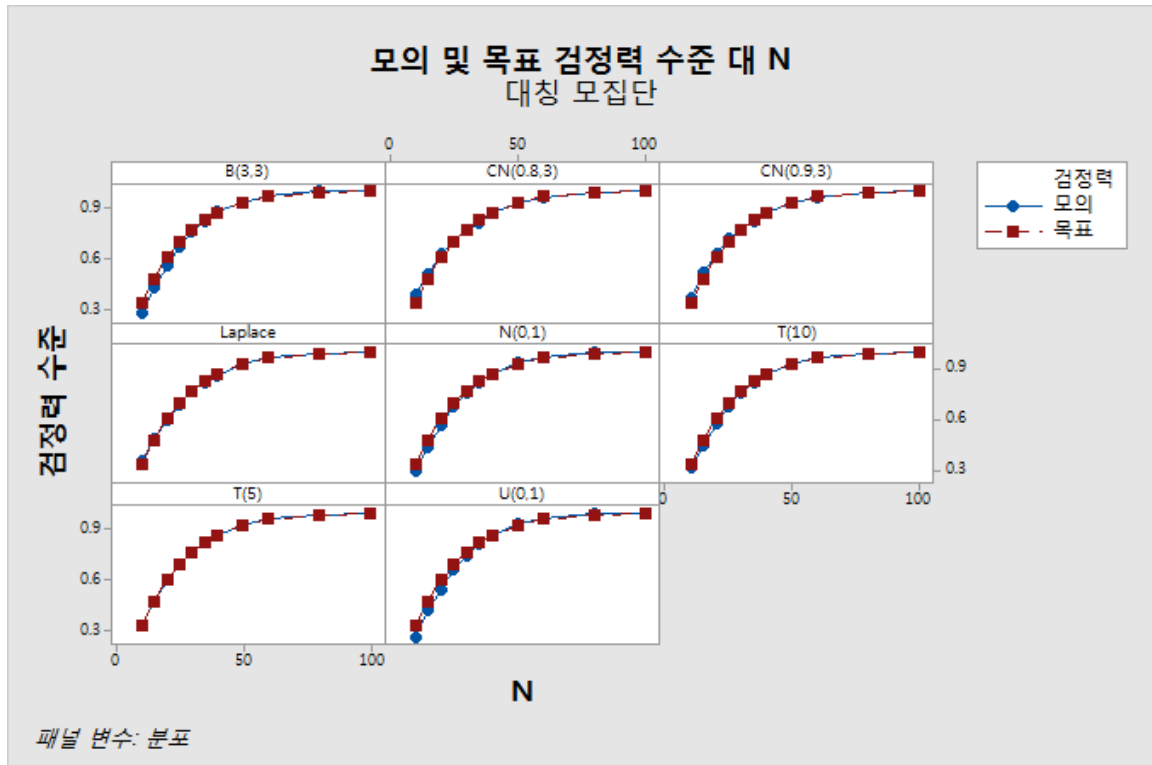


그림 3 표본을 대칭 모집단에서 생성한 경우 $\alpha = 0.05$ 에서 이론적 목표 검정력 곡선과 비교한 양측 1-표본 t-검정의 모의 검정력 곡선. 검정력 값은 차이 $\delta = \sigma/2$ 에서 평가됩니다.

그러나 표본을 치우친 모집단에서 추출할 때는 표 4와 그림 4에 나와 있는 것처럼 표본이 작을 경우 모의 검정력 값이 목표값에서 벗어납니다. 자유도가 5인 카이-제곱 분포와 자유도가 10인 카이-제곱 분포처럼 약간 치우친 모집단의 경우 표본 크기가 20개 이상일 때 목표 및 모의 검정력 수준은 유사합니다. 예를 들어 이면 카이 제곱 5 및 카이 제곱 10 분포에 대한 모의 검정력 수준이 각각 0.576 및 0.577일 때 목표 검정력 수준은 0.565입니다. 극도로 치우친 분포의 경우에는 모의 검정력 수준이 목표 유의 수준에 근접하려면 더 큰 표본이 필요합니다. 그 이유는 표본 크기가 작고 모집단이 극도로 치우친 경우 1-표본 t-검정에서 제1종 오류가 올바르게 관리되지 않기 때문일 수 있습니다.

표 4 표본을 채우친 모집단에서 추출했을 때 차이 $\delta = \sigma/2$ 에서 크기가 $\alpha = 0.05$ 인 양측 1-표본 t-검정의 모의 검정력 값. 모의 검정력 값은 정규 가정 하에서 도출된 목표 검정력 값과 비교됩니다.

N	목표 검정력	지수		Chi(3)	B(8,1)	Chi(5)	Chi(10)
		모집단 왜도					
		2.0		1.633	-1.423	1.265	0.894
		모의 검정력 수준					
10	0.293	0.206		0.212	0.39	0.225	0.238
15	0.438	0.416		0.413	0.484	0.409	0.407
20	0.565	0.604		0.591	0.566	0.576	0.577
25	0.67	0.763		0.734	0.657	0.709	0.695
30	0.754	0.859		0.834	0.729	0.808	0.785
35	0.82	0.917		0.895	0.776	0.874	0.835
40	0.869	0.955		0.935	0.823	0.925	0.905
50	0.934	0.987		0.981	0.90	0.973	0.96
60	0.968	0.997		0.994	0.937	0.991	0.985
80	0.993	1.00		0.999	0.98	0.999	0.997
100	0.999	1.00		1.00	0.994	1.00	1.00

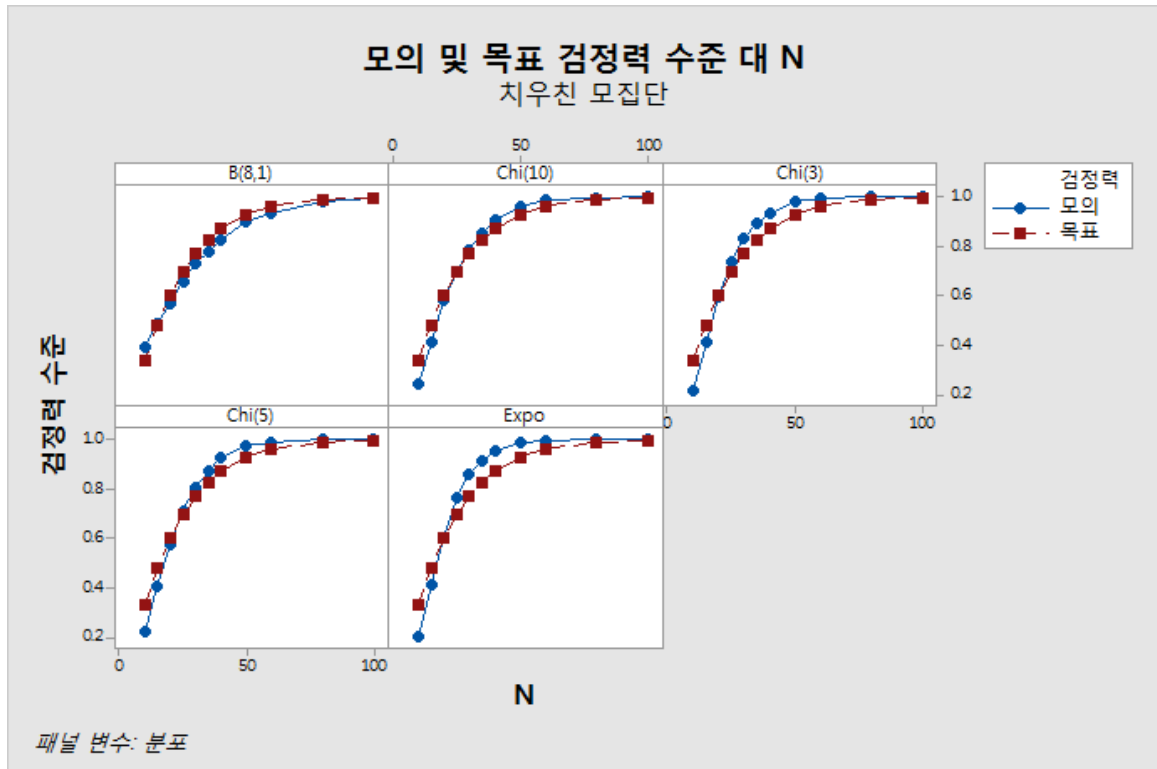


그림 4 표본을 대칭 모집단에서 추출했을 때 $\alpha = 0.05$ 에서 이론적 목표 검정력 곡선과 비교한 양측 1-표본 t-검정의 모의 검정력 곡선. 검정력 값은 차이 $\delta = \sigma/2$ 에서 평가됩니다.

요약하면, 약간 치우친 분포의 경우 검정력 함수는 표본 크기가 20 이상일 때 표본이 추출된 모집단에 관계 없이 신뢰할 수 있습니다. 극도로 치우친 모집단의 경우에는 모의 검정력이 목표 검정력과 가까우려면 더 큰 표본 크기(약 40)가 필요합니다.

© 2015, 2017 Minitab Inc. All rights reserved.
Minitab®, Quality. Analysis. Results.® and the Minitab® logo are all registered trademarks of Minitab, Inc., in the United States and other countries. See minitab.com/legal/trademarks for more information.