本書は、Minitab 統計ソフトウェアのアシスタントで使用される方法およびデータチェックを開発するため、Minitab の統計専門家によって行われた調査に関する一連の文書群を構成する文書の 1 つです。

1 サンプル t 検定

概要

1 サンプル t 検定は、工程の平均を推定し、平均と目標値を比較するために使用されます。この検定は、サンプルが中規模の場合、正規性の仮定に極めて鈍感であるため、頑健な手順と考えられています。ほとんどの統計学の教科書によれば、1 サンプル t 検定と平均の t-信頼区間は、サイズが 30 以上のサンプルに適しています。

本書では、最小サンプル単位が 30 という一般規則を評価するために行ったシミュレーションについて説明します。シミュレーションでは、1 サンプル t 検定に対する非正規性の影響に焦点を当てました。また、検定結果に対する異常なデータの影響も評価する必要がありました。

調査に基づき、アシスタントでは、データに対して次のチェックが自動的に実行され、レポートカードに結果が表示されます。

- 異常なデータ
- 正規性（正規性が問題にならないぐらいにサンプルが十分に大きいか）
- サンプルサイズ

1 サンプル t 検定の手法に関する一般的な情報は、Arnold（1990）、Casella and Berger（1990）、Moore and McCabe（1993）、Srivastava（1958）を参照してください。

注　対応のある t 検定では、1 サンプル t 検定の方法が対応のある差のサンプルに適用されるため、本書の結果はアシスタントの対応のある t 検定にも適用されます。

データチェック

異常なデータ

異常なデータとは、極端に大きいまたは小さいデータ値を指し、外れ値とも呼ばれています。異常なデータは、分析の結果に強い影響を与える可能性があります。サンプルが小さい場合には、統計的に有意な結果を検出する確率に影響することがあります。異常なデータは、データ収集での問題を示している可能性や、分析している工程の異常な動作に起因する場合があります。これらのデータ点は調べる価値があり、可能な場合には修正する必要があります。

目的

分析の結果に影響する可能性がある、全体のサンプルに比べて非常に大きい、または非常に小さいデータ値をチェックする方法を開発する必要がありました。

方法

箱ひげ図の外れ値を特定するために、Hoaglin, Iglewicz, and Tukey (1986) によって説明された方法に基づいて、異常データをチェックする方法を開発しました。

結果

アシスタントでは、分布の下位四分位数を下回る幅または上位四分位数を上回る幅が四分位間範囲の 1.5 倍より大きいデータ点は異常と識別されます。下位四分位数および上位四分位数とは、データの 25 番目および 75 番目の百分位数を指します。四分位間範囲とは、2 つの四分位数間の差を指します。この方法は、特定の各外れ値を検出することが可能なため、複数の外れ値がある場合でも有効に機能します。

異常なデータを調べる場合、1 サンプル t 検定のアシスタントレポートカードには次のステータスインジケータが表示されます。

ステータス	状態
	異常なデータ点はありません。
	少なくとも 1 つのデータ点が異常です。検定結果に影響する可能性があります。

正規性

1 サンプル t 検定は、母集団が正規分布に従っているという仮定のもとで導出されます。データが正規分布に従っていない場合でも、サンプルサイズが十分に大きければ、この検定は有効に機能します。

目的

サンプルサイズおよび正規性のガイドラインを提供するために、検定のタイプ I の誤りとタイプ II の誤りに対する非正規性の効果を判断する必要がありました。

方法

1 サンプル t 検定を実行するとき、または母集団の平均の t-信頼区間を検出するときに正規性の仮定を無視できるサンプルサイズを判断するために、シミュレーションを行いました。

最初の分析は、検定のタイプ I 過誤率に対する非正規性の効果を評価するために設計しました。具体的には、検定が母集団の分布に鈍感であるために必要な最小サンプルサイズを推定する必要がありました。正規分布と非正規分布から生成された小規模、中規模、および大規模サンプルで 1 サンプル t 検定を実行しました。非正規母集団には、軽度から重度に歪んだ母集団、軽い裾から重い裾を持つ対称分布、および混合正規分布が含まれます。正規母集団は、比較のための対照母集団となります。各ケースで、シミュレートした有意水準を計算し、目標（名目）有意水準の 0.05 と比較しました。検定が正しく行われれば、シミュレートした有意水準は 0.05 に近くなります。分布に関係なく、目標水準に常に近くなる最小サンプルサイズを評価するために、すべての異なる条件全体のシミュレートした有意水準を調べました。詳細は、「付録 A」を参照してください。

2 つ目の分析では、検定のタイプ II の誤りに対する非正規性の効果を調べました。シミュレーションの設計は、1 つ目の分析と同一です。ただし、異なる条件でシミュレートした検出力水準と、1 サンプル t 検定の理論上の検出力関数を使用して計算した目標検出力水準を比較しました。詳細は、「付録 B」を参照してください。

結果

検定のタイプ I 過誤率およびタイプ II 過誤率の両方に対する非正規性の効果は、サンプルサイズ 20 で最小限です。ただし、サンプルの親母集団が極端に歪んでいると、より大きなサンプルが必要になる場合があります。このような場合でも、約 40 のサンプルサイズをお勧めします。詳細は、「付録 A」および「付録 B」を参照してください。

検定が比較的小さなサンプルで正しく行われるため、アシスタントではデータの正規性は検定されません。代わりに、サンプルのサイズを調べ、レポートカードに次のステータスインジケータが表示されます。

ステータス	状態
	サンプルサイズが少なくとも 20 です。正規性に問題はありません。
	サンプルサイズが 20 未満です。正規性に問題がある可能性があります。

サンプルサイズ

一般的に、仮説検定は、「差なし」の帰無仮説を棄却する証拠を集めるために実行されます。サンプルが小さすぎると、実際に平均に差があるときに、その差を検出するのに検定の検出力が十分でない場合あり、その結果、タイプ II の誤りが生じます。したがって、サンプルサイズが実質的に重要な差を検出するのに十分な大きさであることを確認することが重要です。

目的

データから帰無仮説に反する十分な証拠が得られない場合、検定のサンプルサイズが、高い確率で対象となる実質的な差を検出するのに十分な大きさかどうかを判断する必要があります。サンプルサイズの計画の目的は、サンプルサイズが、重要な差を高い確率で検出するのに十分な大きさであることを確認することですが、大きすぎて無意味な差が高い確率で統計的に有意になるようではいけません。

方法

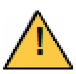
サンプルサイズ分析の検出力は、統計分析を実行するために使用される特定の検定の理論上の検出力関数に基づきます。前述したように、1 サンプル t 検定の検出力関数は、サンプルサイズが少なくとも 20 のとき、正規性の仮定に鈍感です。検出力関数は、サンプルサイズ、目標平均と母集団平均の差、および母集団の分散によって異なります。詳細は、「付録 B」を参照してください。

結果

データから帰無仮説に反する十分な証拠が得られない場合、アシスタントでは与えられたサンプルサイズで 80%および 90%の確率で検出できる実質的な差が計算されます。さらに、ユーザーが対象の特定の実質的な差を指定すると、80%および 90%の確率で差を検出できるサンプルサイズが計算されます。

結果はユーザーの特定のサンプルによって異なるため、報告する一般的な結果はありません。1 サンプル t 検定の検出力についての詳細は、「付録 B」を参照してください。

検出力とサンプルサイズを調べる場合、1 サンプル t 検定のアシスタントレポートカードには次のステータスインジケータが表示されます。

ステータス	状態
	検定で平均と目標の差が検出されます。検出力に問題はありません。 または 検出力は十分です。検定で平均と目標の差は検出されませんでしたが、サンプルは少なくとも 90%の確率で差を検出するのに十分な大きさです。
	検出力は十分な可能性があります。検定で平均と目標の差は検出されませんでしたが、サンプルは 80%～90%の確率で差を検出するのに十分な大きさです。90%の検出力を達成するのに必要なサンプルサイズが報告されます。
	検出力は十分でない可能性があります。検定で平均と目標の差は検出されませんでした。サンプルは 60%～80%の確率で差を検出するのに十分な大きさです。80%および 90%の検出力を達成するのに必要なサンプルサイズが報告されます。

ステータス	状態
	検出力は十分ではありません。検定で平均と目標の差は検出されませんでした。サンプルは少なくとも 60%の確率で差を検出するのに十分な大きさではありません。80%および 90%の検出力を達成するのに必要なサンプルサイズが報告されます。
	検定で平均と目標の差が検出されませんでした。ユーザーが、検出する実質的な平均と目標の差を指定しなかったため、サンプルサイズ、標準偏差および α に基づいて 80%および 90%の確率で検出できる差がレポートに示されます。

参考文献

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., and 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., Iwaszkiewicz, 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.*

付録 A：有意水準（検定の妥当性）に対する非正規性の影響

正規性の仮定では、1 サンプル t 検定は一様最強力（UMP）不偏サイズα検定です。つまりこの検定には、平均に関して、他の不偏サイズα検定と同等以上の性能があります。ただし、サンプルの親母集団が正規分布に従っていない場合、上記の最適性特性は、サンプルサイズが十分に大きい場合に当てはまります。言い換えると、十分に大きいサンプルの場合、1 サンプル t 検定の実際の有意水準は、正規データおよび非正規データの目標水準にほぼ等しく、また検定の検出力関数は正規性の仮定に鈍感です（Srivastava 1958）。

t 検定が正規性の仮定に鈍感であるのに十分な大きさとみなされるサンプルサイズを判断する必要がありました。多くの教科書には、サンプルサイズ$n \geq 30$の場合、ほとんどの実用目的では正規性の仮定を無視できると書かれています（Arnold 1990、Casella & Berger 1990、Moore & McCabe 1993）。以下の付録で説明する調査の目的は、1 サンプル t 検定に対するさまざまな非正規分布の影響を調べることによって、この一般規則を評価するシミュレーションの分析を行うことです。

シミュレーションの分析 A

分布に関係なく、安定性があり、目標過誤率に常に近くなる最小サンプルサイズを評価するため、検定のタイプ I 過誤率に対する非正規性の影響を調べる必要がありました。

そのために、異なる特定を持つ複数の分布から生成されたさまざまなサイズ（$n = 10, 15, 20,\ 25,\ 30, 35, 40, 50, 60, 80, 100$）のランダムサンプルを使用して両側 t 検定が実行されました。これらの分布には、以下が含まれます。

- 標準正規分布（N(0, 1)）
- 対称的で裾の重い分布。自由度 5 と 10 の t 分布（t(5)、t(10)）など。
- 位置 0、尺度 1 のラプラス分布（Lpl）
- 歪んだ裾の重い分布。尺度 1 の指数分布（Exp）、自由度 3、5、10 のカイ二乗分布（Chi(3)、Chi(5)、Chi(10)）など。
- 対称的で裾の軽い分布。一様分布（U(0, 1)）および 2 つのパラメータを 3 に設定したベータ分布（B(3, 3)）など。
- 左方向に歪んだ裾の重い分布（B(8, 1)）

さらに、外れ値の直接効果を評価するために、次のように定義される混合正規分布からサンプルを生成しました。

$$\mathrm{CN}\big(\mathrm{p}, \sigma\big) = \mathrm{pN}(0{,}1) + (1-\mathrm{p})\mathrm{N}(0, \sigma)$$

ここで、pは混合パラメータとして定義され、$1-\mathrm{p}$は混入の比率（外れ値の比率）です。この分析のために、$\mathrm{CN}(0.9{,}3)$（母集団メンバーの 10%が外れ値）、$\mathrm{CN}(.8{,}3)$（母集団メンバーの 20%が外れ値）という 2 つの混合正規母集団を選択しました。これら 2 つの分布は対称的で、外れ値に起因する長い裾があります。

各サンプルサイズで、各母集団から 10,000 個のサンプル反復が抽出され、各 10,000 サンプルについて帰無仮説$\mu = \mu_o$および対立仮説 $\mu \neq \mu_o$を使用した 1 サンプル t 検定が実行されました。各検定で、仮説の平均μ_oをサンプルの親母集団の真の平均に設定しました。その結果、与えられたサンプルサイズでサンプル反復数 10,000 のうち帰無仮説が棄却される割合が、検定のシミュレートしたタイプ I 過誤率または有意水準を表します。目標有意水準は 5%で、シミュレーション誤差は約 0.2%です。

シミュレーションの結果を表 1 と表 2 に示し、グラフを図 1 と図 2 に示します。

結果と要約

結果（表 1 と図 1 を参照）は、サンプルが対称的な母集団から生成されると、サンプルサイズが小さい場合でも、検定のシミュレートした有意水準は目標有意水準に近くなります。ただし、裾の重い対称的な分布の場合は、混合正規分布から生成される小さいサンプルも含め、サンプルサイズが小さいと、検定結果がやや保守的になります。また、サンプルが小さいと、外れ値によって検定の有意水準が低下するように見えます。ただし、対称的で裾がより軽い親母集団（ベータ分布 Beta(3,3)および一様分布）から小さいサンプルが生成される場合は、この効果は逆になります。これらのケースでは、シミュレートした有意水準は少し高めになります。

表 1　対称的な母集団から生成されたサンプルの両側 1 サンプル t 検定のシミュレートした有意水準。目標有意水準は$\alpha = 0.05$です。

分布	N(0,1)	t(5)	t(10)	Lpl	CN(.9,3)	CN(.8,3)	B(3,3)	U(0,1)
N	対称的で重い裾						対称的でより軽い裾	
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

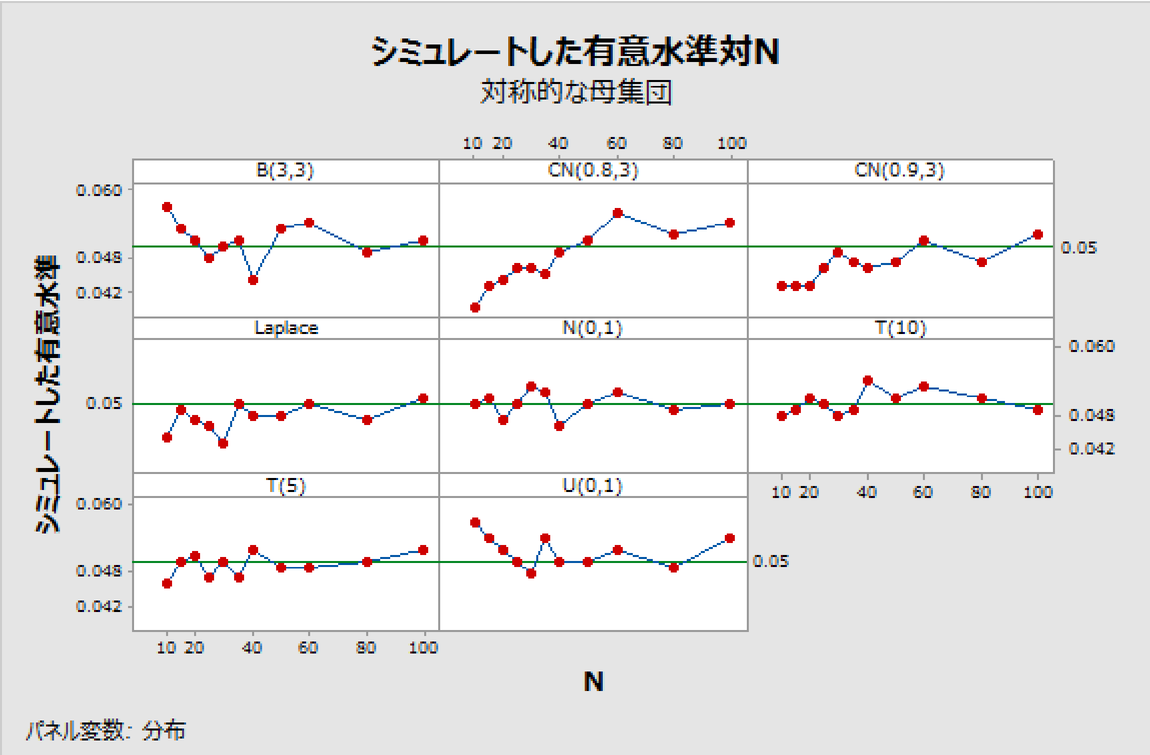


図 1　両側 1 サンプル t 検定のシミュレートした有意水準と対照的な母集団から生成されたサンプルのサイズを対比したプロット。目標有意水準は $\alpha = 0.05$ です。

一方、サンプルが歪んだ分布から生成される場合、検定の性能は歪みの重大度によって異なります。表 2 と図 2 の結果は、1 サンプル t 検定が小さいサンプルでは歪みに敏感であることを示しています。かなり偏った母集団（指数、Chi(3)、Beta(8,1)）の場合、シミュレートした有意水準が目標有意水準近くになるには、より大きなサンプルが必要です。ただし、中程度に歪んだ母集団（Chi(5) と Chi(10)）の場合、最小サンプルサイズの 20 はシミュレートした有意水準が目標水準に近くなるのに十分です。サンプルサイズに 20 を使用すると、シミュレートした有意水準は、自由度 5 のカイ二乗分布で約 0.063、自由度 10 のカイ二乗分布で約 0.056 になります。

表 2　歪んだ母集団から生成されたサンプルの両側 1 サンプル t 検定のシミュレートした有意水準。目標有意水準は $\alpha = 0.05$ です。

N	Exp	Chi(3)	B(8,1)	Chi(5)	Chi(10)
	母集団の歪み				
	2.0	1.633	-1.423	1.265	0.894
	シミュレートした有意水準				
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

N	Exp	Chi (3)	B (8, 1)	Chi (5)	Chi (10)
	母集団の歪み				
	2.0	1.633	-1.423	1.265	0.894
	シミュレートした有意水準				
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

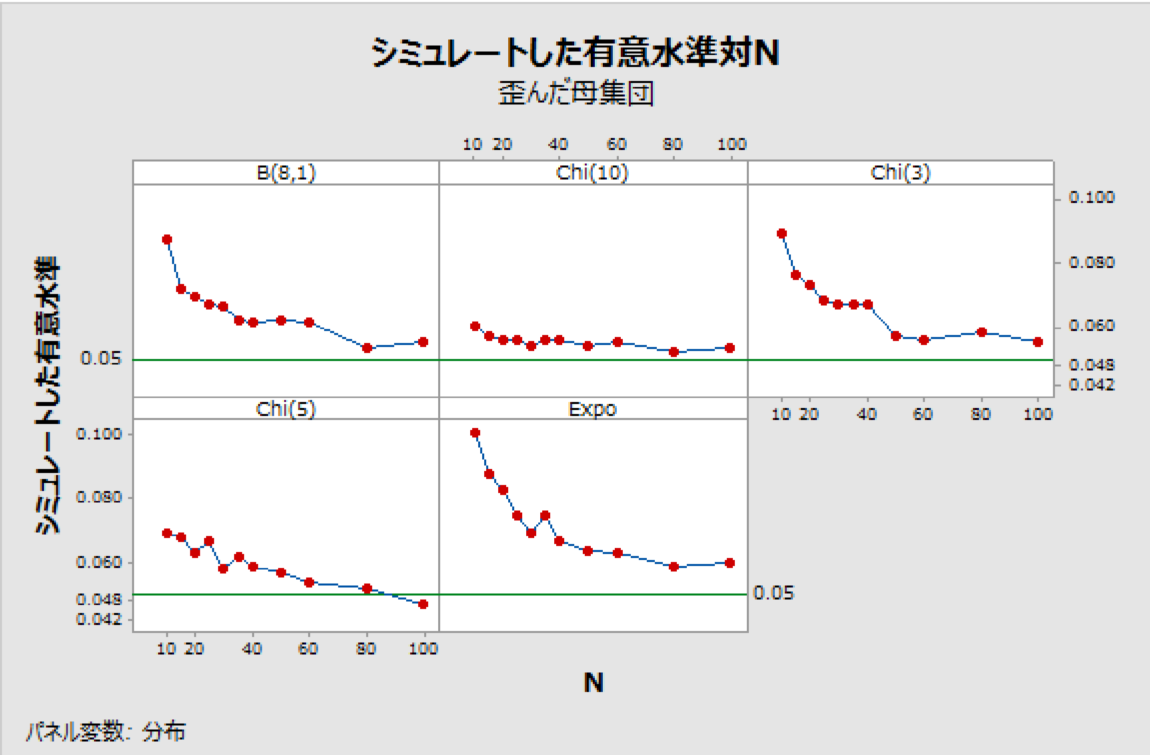


図 2　両側 1 サンプル t 検定のシミュレートした有意水準と歪んだ母集団から生成されたサンプルのサイズを対比したプロット。目標有意水準は$\alpha = 0.05$です。

この調査では、信頼区間ではなく、仮説検定に焦点を当てました。ただし、仮説検定と信頼区間は、どちらも統計的有意性を判断するために使用されるので、結果は必然的に信頼区間にまで拡大適用されます。

付録 B：検定のサンプルサイズと検出力

検出力関数が導出される正規性の仮定に対する検出力関数の感度を調べる必要がありました。βが検定のタイプ II の誤りである場合、$1-\beta$は検定の検出力です。その結果、計画サンプルサイズは、タイプ II 過誤率が小さく（または検出力水準が適度に高く）なるように判断されます。

t 検定の検出力関数は、十分に証明された既知の関数です。Pearson and Hartley（1954）、Neyman, Iwaszkiewicz, and Kolodziejczyk（1935）によって、検出力関数の図表が提供されています。

サイズαの両側 1 サンプル t 検定の場合、この関数の数式は、サンプルサイズおよび真の平均μと仮説の平均μ_oの差に関して、次のように与えられることがあります。

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

ここで、$\mathrm{F}_{\mathrm{d},\lambda}(.)$は自由度$\mathrm{d}=\mathrm{n}-1$と非心パラメータ

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

を使用した非心 t 分布の CDF です。ここで、$\mathrm{t}_{\mathrm{d}}^{\alpha}$は自由度$\mathrm{d}$の t 分布の 100$\alpha$上位百分位数を示します。

片側の対立仮説の場合、

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

として、$\mu > \mu_o$に対する帰無仮説の検定の検出力として与えられ、

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

として、$\mu < \mu_o$に対する帰無仮説の検定の検出力が与えられます。

これらの関数は、データが正規分布に従っており、検定の有意水準はある値αに固定されているという仮定のもとで導出されます。

シミュレーションの分析 B

このシュミレーションは、1 サンプル t 検定の理論上の検出力関数に対する非正規性の効果を評価するために設計しました。非正規性の効果を評価するために、シミュレートした検出力水準と、検定の理論上の検出力関数を使用して計算された目標検出力水準を比較しました。

1 番目のシミュレーションの分析（「付録 A」を参照）で前述した同じ母集団から生成されたさまざまなサイズ（$\mathrm{n} = 10, 15, 20, 25, 30, 35, 40, 50, 60, 80, 100$）のランダムサンプルで$\alpha = 0.05$を使用して両側 t 検定を実行しました。

各母集団の検定の帰無仮説は$\mu = \mu_o - \delta$、対立仮説は$\neq \mu_o - \delta$になります。ここで、μ_oは母集団の真の平均で設定され、$\delta = \sigma/2$（σは親母集団の標準偏差）となります。し

たがって、真の平均と仮説の平均の差は 0 になるため、適切な判断は、帰無仮説を棄却することです。

与えられた各サンプルサイズで、各分布から 10,000 個のサンプル反復が抽出されます。与えられた各サンプルサイズで反復数 10,000 のうち帰無仮説が棄却される割合が、与えられたサンプルサイズと差 δ で行われる検定のシミュレートした検出力水準を表します。サンプルサイズが小さい場合に比較的小さい検出力が得られるため、この特定の差の値を選択しました。

さらに、シミュレートした検出力と比較するために、目標検出力と呼ばれる対応する理論上の検出力が、差 δ とさまざまなサンプルサイズで計算されます。

シミュレーションの結果を表 3 と表 4 に、グラフを図 3 と図 4 に示します。

結果と要約

結果から、サンプルサイズが十分に大きい場合、1 サンプル t 検定の検出力は正規性の仮定に対し一般に鈍感であることが確認できます。対称的な母集団から生成されたサンプルでは、表 3 の結果は、小さいサンプルでも目標有意水準とシミュレートした有意水準が近くなることを示しています。図 3 に示した対応する検出力曲線は、実質的に区別できません。混合正規分布から生成されたサンプルでは、検出力は小規模から中規模のサンプルサイズでやや保守的になります。これは、これらの母集団では検定の実際の有意水準が、固定された目標有意水準αより少し高くなることに起因している可能性があります。

表 3　サンプルが対称的な母集団から生成される場合に、サイズ$\alpha = 0.05$の両側 1 サンプル t 検定で差$\delta = \sigma/2$でシミュレートした検出力水準。シミュレートした検出力水準は、正規性の仮定のもとで導出された理論上の目標検出力水準と比較されます。

n	目標検出力	N(0,1)	t(5)	t(10)	Lpl	CN(.9,3)	CN(.8,3)	B(3,3)	U(0,1)
		$\boldsymbol{\delta = \sigma/2}$（対称的な母集団）でシミュレートした検出力水準							
10	0.293	0.299	0.334	0.311	0.357	0.361	0.385	0.280	0.269
15	0.438	0.438	0.480	0.450	0.491	0.512	0.511	0.423	0.421
20	0.565	0.570	0.603	0.578	0.600	0.629	0.623	0.557	0.548
25	0.670	0.674	0.695	0.680	0.691	0.712	0.700	0.665	0.670
30	0.754	0.756	0.770	0.756	0.767	0.768	0.765	0.754	0.750
35	0.820	0.819	0.827	0.815	0.820	0.819	0.812	0.822	0.818
40	0.869	0.870	0.871	0.868	0.862	0.869	0.868	0.875	0.867
50	0.934	0.933	0.929	0.930	0.929	0.923	0.925	0.932	0.940
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

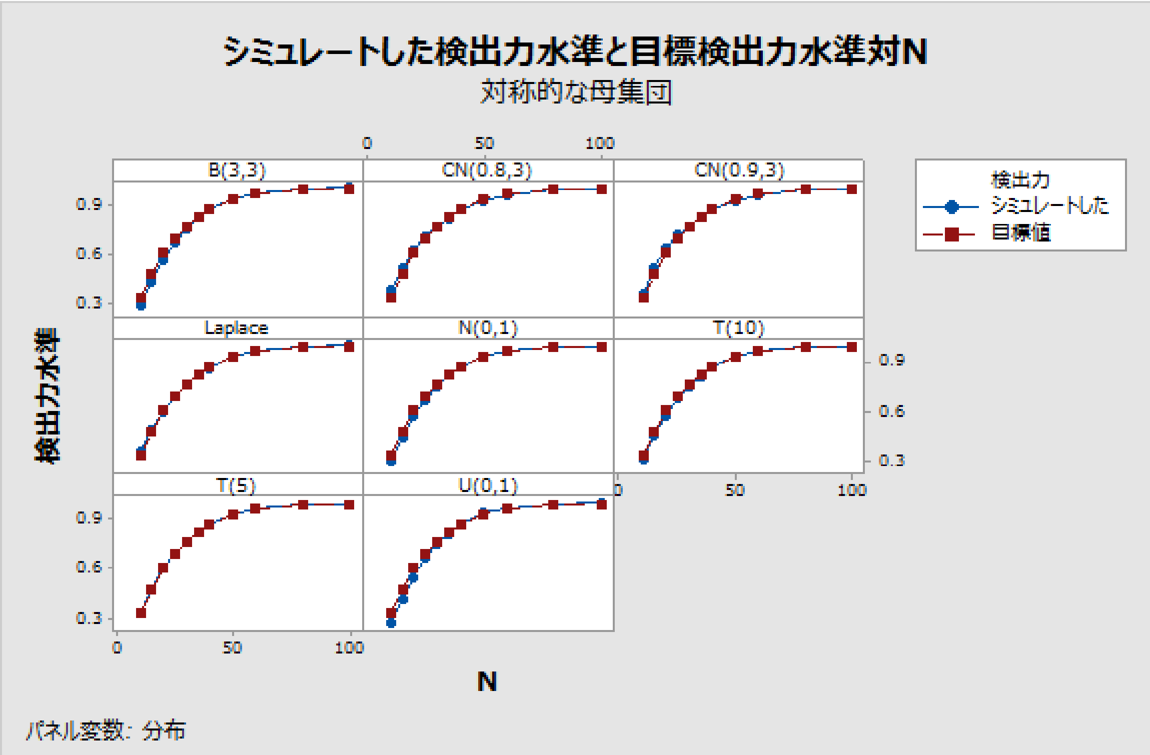


図 3　サンプルが対称的な母集団から生成される場合に、両側 1 サンプル t 検定で $\alpha = 0.05$ でシミュレートした検出力曲線と理論上の目標検出力曲線の比較。検出力は、 $\delta = \sigma/2$の差で評価されます。

ただし、サンプルが歪んだ母集団から抽出される場合、表 4 と図 4 に示すように、シミュレートした検出力は目標から外れます。自由度 5 のカイ二乗分布や自由度 10 のカイ二乗分などの中程度に歪んだ母集団の場合、サンプルサイズが少なくとも 20 のとき、目標検出力水準とシミュレートした検出力水準は近くなります。たとえば、$n = 20$では、目標検出力水準は 0.565 で、シミュレートした検出力はカイ二乗分布 Chi(5) と Chi(10) でそれぞれ 0.576、0.577 となります。極端に歪んだ母集団の場合、シミュレートした検出力水準が目標有意水準に近づくには、より大きなサンプルが必要です。これは、サンプルサイズが小さく、親母集団が極端に歪んでいる場合、1 サンプル t 検定でタイプ I の誤りが適切に制御されないことに起因している可能性があります。

表 4　サンプルが歪んだ母集団から生成される場合に、サイズ$\alpha = 0.05$の両側 1 サンプル t 検定で差$\delta = \sigma/2$でシミュレートした検出力水準。シミュレートした検出力水準は、正規性の仮定のもとで導出された理論上の目標検出力と比較されます。

		Exp		Chi(3)	B(8,1)	Chi(5)	Chi(10)
			母集団の歪み				
		2.0		1.633	-1.423	1.265	0.894
N	目標検出力		シミュレートした検出力水準				
10	0.293	0.206		0.212	0.390	0.225	0.238

		Exp		Chi(3)	B(8,1)	Chi(5)	Chi(10)
			母集団の歪み				
		2.0		1.633	-1.423	1.265	0.894
N	目標検出力		シミュレートした検出力水準				
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	0.670	0.763		0.734	0.657	0.709	0.695
30	0.754	0.859		0.834	0.729	0.808	0.785
35	0.820	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	0.900	0.973	0.960
60	0.968	0.997		0.994	0.937	0.991	0.985
80	0.993	1.000		0.999	0.980	0.999	0.997
100	0.999	1.000		1.000	0.994	1.000	1.000

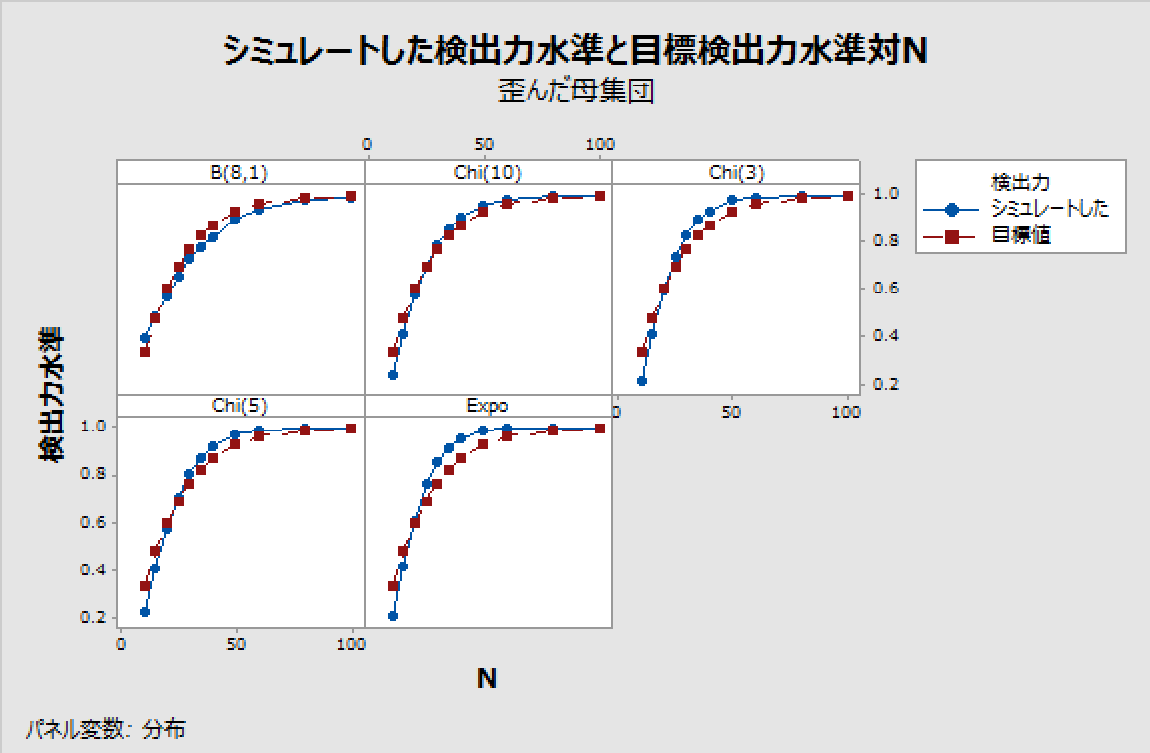


図 4　サンプルが対称的な母集団から生成される場合に、両側 1 サンプル t 検定で $\alpha = 0.05$ でシミュレートした検出力曲線と理論上の目標検出力曲線の比較。検出力は、 $\delta = \sigma/2$ の差で評価されます。

要約すると、中程度に歪んだ分布では、サンプルが抽出される親母集団に関係なく、サンプルサイズが少なくとも 20 のとき、検出力関数は信頼できます。極端に歪んだ母集団の場合、シミュレートした検出力が目標検出力に近くなるには、より大きなサンプル（約 40）が必要です。