サンプル数の理論的決め方

基本事項

・ここでは母集団の分布は正規分布を仮定して説明しています。

・**母集団の分布の形にかかわらず**、平均μ、分散σ^2の母集団からサンプリングしたサンプル平均値$\sum_{i=1}^{n} x_i \big/ n$の分布は、サンプル数$n$が多くなると平均$\mu$、分散$\sigma^2/n$の正規分布に近づきます（中心極限定理）。
例えば、１個の試料を反復測定した平均値の場合、反復数が多くなると各試料の平均値の分布は正規分布に近づきます。

・１個の試料の反復測定回数が少ない場合、その測定平均値の分布が理論的に正規分布するとは言い切れませんので分布図を作成して分布の形を確認することが大切です。

・検討対象の分布が正規分布とは言えない場合は、変数変換により正規分布に近づけることができる場合があります。

・例えば、測定値xの分布が右側に長く裾を引く分布の場合、対数変換した$\log(x)$の分布は左右対称に近づき、正規分布に近くなります。

・検討対象の分布が正規分布かどうかの確認方法には、正規確率プロットのような図を用いる方法、尖度（とがりど）、歪度（わいど、ひずみど）という統計量の検定により判断する方法、Shapiro-Wilk 検定により判断する方法などがあります。

・サンプル平均値は母集団平均値の推定値のため、サンプル平均値の信頼区間は母集団平均値の存在する推定範囲を示しています。サンプル数nが多くなるほど母集団平均値の推定範囲が狭くなります。

・**母集団の分布の形やサンプル数ｎの大きさにかかわらず**、分散がσ^2の母集団からn個サンプリングしたサンプル平均値の分散はσ^2/nになります。

・サンプルの不偏標準偏差$s_{(n-1)}^2$は母集団の分散σ^2の推定値になります。

サンプル数nが多くなるほど$s_{(n-1)}^2$はσ^2に近づきます。

・比率は目的とする事象が当てはまるサンプルを１、当てはまらないサンプルを０とした場合の平均値に等しくなります。

・比率の分布として重要な２項分布を正規分布で近似可能な条件には経験則として３通りあります。

1) np と nq の小さい方が 10 より大きい。10 の代わりに 5 が用いられる場合あります。
2) $0.1 \le p \le 0.9$ で、かつ $5 < npq$ 。
3) $25 < npq$

ここで、 n はサンプル数、 p はサンプル内の比率、 $q = 1 - p$ です。

参考文献：蓑谷千夫彦凰彦 (2003). 統計分布ハンドブック，朝倉出版，東京，p.483.

・母集団の比率 P の分散は $P(1-P)$、この母集団から n 個サンプリングしたサンプル中の比率の期待値は nP、分散は $P(1-P)/n$ になります。

注）n が非常に大きいとき以外は、このサンプルの分散 $P(1-P)/n$ を用いて信頼区間を計算しても実用的な精度がないと言われています。
信頼区間の計算には Clopper&Pearson(1934)の信頼区間をはじめ、近似精度によっていくつかの計算方法が提案されています。

>>　計算例 PDF(141KB、A4 サイズ 3 ページ)

参考文献：竹内啓、藤野和健 (1981). 2 項分布とポアソン分布，東京大学出版会，東京，pp.158-164.

1.　母集団の推定値が平均値の場合

$$n \approx \left(\frac{2k}{CI}\right)^2 s^2 \tag{1.1}$$

ここで、n は理論的サンプル数、CI（Confidence Interval）は信頼区間、k は 1.96（危険率 5%のときの正規分布の棄却限界）または 2.58（危険率 1%のときの正規分布の棄却限界）、s^2 は母集団の分散の推定値。この式は母集団の大きさ N がサンプル数 n よりはるかに大きいときの近似式です。

上記の式でサンプル数 n を計算するためには母集団の分散の推定値が必要です。この推定値は日頃の品質管理データや予備調査データがあれば、そのデータの不偏分散を計算して代入します。データがない場合は予備調査を行うか、文献値を探すか、または化学分析値の場合は Horwitz の式を用いて計算した分散が使えないか検討します。化学分析値の分散の大きさは分析値の濃度が低くなるにつれ大きくなりますので、調査目的に合った濃度（例えば基準値付近の

濃度）の分散を代入します。

例題

殻なし落花生のアフラトキシン検査で、ロットから 5kg 抜き取ったサンプルを粉砕し、粉砕したサンプル 100g を薄層クロマトグラフィ（TLC）で 1 回分析した場合、アフラトキシン濃度 15ppb(B_1,B_2,G_1,G_2 の合計濃度)のときの分析値の分散の大きさは 491 です。同じ条件で 20kg サンプリングした場合の分散は 211 です。これらの分散の値は以下の文献内の分析値のばらつきに関するモデル式を用いて計算した値で、ロットのアフラトキシン濃度の不偏分散です。

FAO (1993). Sampling plans for aflatoxin analysis in peanuts and corn, FAO, Rome, p.39.

Whitaker, T.B., Dickens, J.W. and Monroe, R.J. (1974). Variability of aflatoxin test results. J. Am. Oil Chem. Soc., 51, 214-218.

Codex のアフラトキシンの基準値 15ppb(total aflatoxin)において、15ppb の±20%の±3ppb の精度でロットのアフラトキシン濃度を 5kg のサンプルから知りたいとします。このときの危険率を 5%で考える場合は、±3ppb の範囲は 95%信頼区間として考えることになり k =1.96、CI =3×2=6（ppb）、s^2 =491 を上記の式に代入して、

$$n \approx \left(\frac{2\times 1.96}{6}\right)^2 \times 491 = 209.58 \approx 210 \qquad (1.2)$$

が得られます。１ロットから 5kg のサンプルを 210 個ランダムサンプリングして分析すれば 15ppb 付近のロットのアフラトキシン濃度を危険率 5%で±3ppb の精度（95%信頼区間）で知ることができます。

15ppb において危険率 1%で±5ppb の精度でロットのアフラトキシン濃度を 20kg のサンプルから知りたい場合は k =2.58、CI =5×2=10（ppb）、s^2 =211 を上記の式に代入して、

$$n \approx \left(\frac{2\times 2.58}{10}\right)^2 \times 211 = 56.18 \approx 57 \qquad (1.3)$$

が得られます。１ロットから 20kg のサンプルを 57 個ランダムサンプリングして分析すれば 15ppb 付近のロットのアフラトキシン濃度を危険率１%で±10ppb の精度（99%信頼区間）で知ることができます。

補足１　理論的サンプル数ｎの計算式の導出

母集団から n 個（ただし、$n \ll$ 母集団の大きさ N ）抜き取ったサンプルの分析

値を x_i （i =1,･･･ n ）とすると母集団の平均値 μ は、

$$\bar{x} - k\sqrt{\frac{s^2}{n}} < \mu < \bar{x} + k\sqrt{\frac{s^2}{n}} \qquad (1.4)$$

の信頼区間内に存在すると推定できます。ここで $\bar{x}$ は n 個の分析値 x_i の平均値、

k は正規分布の棄却限界値、s^2 は母集団の分散の推定値（n 個の分析値 x_i の不偏

分散）です。
信頼区間を CI とすると、CI =信頼上限－信頼下限なので

$$CI = k\sqrt{\frac{s^2}{n}} - \left(-k\sqrt{\frac{s^2}{n}}\right) = 2k\sqrt{\frac{s^2}{n}} \qquad (1.5)$$

となります。この式を n について解くと(1.1)式の

$$n = \left(\frac{2k}{CI}\right)^2 s^2 \qquad (1.6)$$

が得られます。

補足２　（サンプルの大きさｎ）≪（母集団の大きさ N）でない場合
母集団の平均値 μ は、

$$\bar{x} - k\sqrt{\frac{N-n}{N-1}\frac{s^2}{n}} < \mu < \bar{x} + k\sqrt{\frac{N-n}{N-1}\frac{s^2}{n}} \qquad (1.7)$$

の信頼区間内に存在すると推定できます。ここで、$\dfrac{N-n}{N-1}$ は有限修正と呼ばれる

項です。
補足１の近似式同様に n について解くと、

$$n = \frac{N}{\left(\frac{CI}{2k}\right)^2 \frac{N-1}{s^2} + 1} \qquad (1.8)$$

が得られます。
$n \ll N$ の場合、有限修正の項は、

$$\frac{N-n}{N-1} = \frac{1-n/N}{1-1/N} \approx \frac{1}{1} = 1 \qquad (1.9)$$

となり上記(1.6)の近似式が得られます。

補足３　近似式(1.1)式と有限修正のある(1.8)式の比較例

殻なし落花生のアフラトキシン検査に関する上記の例題でロットの大きさを 100t とします。ロットから 20kg 抜き取る場合、N =100000kg/20kg=5000 になります。$n \approx 57$ の(1.3)式と同じ条件を(1.8)式に代入すると、

$$n = \frac{5000}{\left(\frac{10}{2 \times 2.58}\right)^2 \frac{5000-1}{211} + 1} = 55.56 \approx 56 \qquad (1.10)$$

となり、ほぼ等しい結果が得られます。

ロットの大きさが 1t の場合は N =1000kg/20kg=50 となり、

$$n = \frac{50}{\left(\frac{10}{2 \times 2.58}\right)^2 \frac{50-1}{211} + 1} = 26.75 \approx 27 \qquad (1.11)$$

となります。このように母集団の大きさ N がサンプルの大きさ n に比べ十分に大きくない場合は有限修正のある(1.8)式と(1.1)式の近似式の理論的サンプル数はかなり変わります。

また、母集団の大きさ N がサンプルの大きさ n の 100 倍以上あれば近似式と有限修正のある式の理論的サンプル数は近い値になります。

2．　母集団の推定値が比率の場合

$$n \approx \left(\frac{2k}{CI}\right)^2 \hat{P}\left(1-\hat{P}\right) \qquad (2.1)$$

ここで、n は理論的サンプル数、CI（Confidence Interval）は信頼区間、k は 1.96（危険率 5%のときの正規分布の棄却限界）または 2.58（危険率 1%のときの正規分布の棄却限界）、$\hat{P}$ は母集団の比率の推定値で、$\hat{P}\left(1-\hat{P}\right)$ は母集団の分散の推定値です。この式は母集団の大きさ N がサンプル数 n よりはるかに大きいときの近似式です。

理論的サンプル数 n を計算するためには調査で知りたい母集団の比率の推定値 $\hat{P}$ が必要です。$\hat{P}$ は日頃の品質管理データや予備調査データがあれば、そのデータから計算します。データがない場合は予備調査を行うか、文献値を探すなどします。

例題

ある母集団について予備調査を行った結果、基準値を超える割合が 5%だったとします。この母集団について基準値を超える割合を±2%の精度で知りたいとし

ます。このときの危険率を 5%で考える場合は、±2%の範囲は 95%信頼区間になり k =1.96、CI =0.02×2=0.04（4%）、$\hat{P}$ =0.05(5%)を上記の式に代入すると、

$$n \approx \left(\frac{2 \times 1.96}{0.04} \right)^2 \times 0.05 \times (1 - 0.05) = 456 \qquad (2.2)$$

が得られます。この母集団から 456 個をランダムサンプリングすれば基準値を超える割合を危険率 5%で±2%の精度で知ることができます。

補足１　理論的サンプル数ｎの計算式の導出

母集団から n 個（ただし、$n \ll$ 母集団の大きさ N ）抜き取ったサンプル内の比率を $\hat{P}$ とすると、母集団の比率 P は、

$$\hat{P} - k\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} < P < \hat{P} + k\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \qquad (2.3)$$

の信頼区間内に存在すると推定できます。ここで、k は正規分布の棄却限界値、$\hat{P}(1-\hat{P})$ は母集団の分散の推定値です。

信頼区間を CI とすると、CI =信頼上限－信頼下限なので

$$CI = k\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} - \left(-k\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \right) = 2k\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \qquad (2.4)$$

となります。この式を n について解くと(2.1)式の

$$n = \left(\frac{2k}{CI} \right)^2 \hat{P}(1-\hat{P}) \qquad (2.5)$$

が得られます。

補足２　（サンプルの大きさｎ）≪（母集団の大きさ N）でない場合

母集団の比率 P は、

$$\hat{P} - k\sqrt{\frac{N-n}{N-1}\frac{\hat{P}(1-\hat{P})}{n}} < P < \hat{P} + k\sqrt{\frac{N-n}{N-1}\frac{\hat{P}(1-\hat{P})}{n}} \qquad (2.6)$$

の信頼区間内に存在すると推定できます。ここで、$\frac{N-n}{N-1}$ は有限修正と呼ばれる項です。
補足１の近似式同様に n について解くと、

$$n=\frac{N}{\left(\frac{CI}{2k}\right)^2\frac{N-1}{\hat{P}\left(1-\hat{P}\right)}+1} \tag{2.7}$$

が得られます。

$n \ll N$ の場合、有限修正の項は、

$$\frac{N-n}{N-1}=\frac{1-n/N}{1-1/N}\approx\frac{1}{1}=1 \tag{2.8}$$

となり、上記(2.5)の近似式が得られます。

補足３　近似式(2.1)式と有限修正のある(2.7)式の比較

上記の例題で母集団の大きさを 50000 検体とし、$n \approx 456$ の(2.2)式と同じ条件を(2.7)式に代入すると、

$$n=\frac{50000}{\left(\frac{0.04}{2\times1.96}\right)^2\frac{50000-1}{0.05\times\left(1-0.05\right)}+1}=452 \tag{2.9}$$

となり、ほぼ等しい結果が得られます。

母集団の大きさを 5000 検体とした場合は、

$$n=\frac{5000}{\left(\frac{0.04}{2\times1.96}\right)^2\frac{5000-1}{0.05\times\left(1-0.05\right)}+1}=418 \tag{2.10}$$

となります。有限修正のある(2.7)式よりも近似式(2.1)式の方が理論的サンプル数が多いので近似式は安全サイドの見積もりになります。ただし、母集団の大きさ N がサンプルの大きさ n に比べ十分に大きくない場合は有限修正のある(1.8)式と(1.1)式の近似式の理論的サンプル数はかなり変わります。

また、母集団の大きさ N がサンプルの大きさ n の 100 倍以上あれば近似式と有限修正のある式の理論的サンプル数は近い値になります。