

る可能性も考えられる。つまり、ELISA法の基本的な性能として、精度良く分析可能な物質の最低濃度に近い濃度の分析を試みた事による併行精度への影響が考えられる。同じくトワイズミのロットを対象としたサンプル間のばらつきは25.6%と推定され、27.7%と推定された総標準偏差に対して支配的であると考えられる。サンプル間のばらつきが総標準偏差に対して支配的であることは、他のロットについても同様である。具体的には、大麦の総標準偏差が33.8%であるのに対しサンプル間のばらつきは33.4%、ミナミノカオリの総標準偏差が20.0%であるのに対しサンプルのばらつきは18.7%、農林61号ロットの総標準偏差が18.7%であるの対しのばらつきは17.7%である。つまり、ロットによらず、デオキシニバレノール濃度の測定値のばらつきの主たる要因は、試料間での濃度の変動であると言える。

・不確かさの推定

・食肉加工食品中のアセスルファムKあるいは亜硝酸ナトリウム濃度の不確かさの推定

分散分析により、測定値に含まれるばらつきの要因を分析の併行精度

と試料間での化学物質濃度の変動に分離し、それぞれの大きさを推定した。推定された試料間での化学物質濃度の変動を、1製造ロット中の化学物質濃度分布の標準偏差(母標準偏差)として仮定した。また、食品と化学物質との組み合わせごとに得た $n=20$ の試験室代表値の平均値を、1製造ロット中の化学物質濃度の平均(母平均)と仮定した。サンプリングの不確かさ推定の際には、分析精度は一定として不確かさに含めるものとし、サンプルサイズを3、5、10、16と変更した場合の影響についても検討した。

まず、食肉製品と食品添加物との組み合わせによらず、サンプルサイズの増加に伴いサンプリングの不確かさ(SDあるいはRSD)は小さくなった。これは中心極限定理に従い、母平均推定値の分散がサンプルサイズの増加に伴って小さくなるためである。また、サンプルサイズが3の場合でも、ソーセージ中のアセスルファムK並びに亜硝酸ナトリウム、及びハム中のアセスルファムK濃度のサンプリングの不確かさ(RSD)は7.6%が最大であり、これはH21昨年度に報告したサンプルサイズを3と

仮定して推定される生鮮野菜の成分である亜硝酸塩濃度のサンプリングの不確かさに比べてもより小さな値である。また、上記食肉加工品中での各食品添加物の濃度は、生鮮野菜中に含まれる亜硝酸塩濃度の 50～100 分の 1 程度である。母平均と母分散は独立しているが、サンプリングの不確かさ推定値に一定の寄与を考慮している分析精度は、濃度の関数となることが経験的に示されている。このことを考慮しても、製造管理された加工食品中に意図して添加される化学物質濃度の分散は小さいといえる。さらに、規格値等への適合判定を想定し、サンプリングの不確かさが判定の正しさにどの程度の影響を与えるかについて考察するため、サンプリングの不確かさを考慮した場合に母平均が存在すると推定される範囲を、最小と最大の濃度により示した。ソーセージ中のアセスルファミン濃度の母平均が 8.62 ng/g であるのに対し、母平均が存在すると推定される範囲は 8.0～9.3 ng/g、同じくソーセージ中の亜硝酸ナトリウム濃度の母平均が 20.58 ng/g であるのに対し、母平均が存在すると推定される範囲は 19.0～22.1 ng/g であ

る。また、ハム中のアセスルファミン濃度の母平均が 17.27 ng/g であるのに対し、母平均が存在すると推定される範囲は 16.0～18.6 ng/g である。いずれも併行分析数を 3 とした場合の推定であるが、母平均と推定範囲の乖離は小さく、判定を誤る頻度は低いものと想像される。このように、適合判定される場合を想定しても、正しく製造管理されている加工食品については、極めて微量に残留する農薬等、天然汚染物質また季節変動する食品成分などに比べ、より小さなサンプルサイズであっても、十分な確度で判定可能な分析結果が得られることが示唆される。一方で、ハム中の亜硝酸ナトリウムについては、他の食肉加工品中の食品添加物と同程度までサンプリングの不確かさを小さくするためには、サンプルサイズを 16 程度まで大きくする必要がある。また、サンプルサイズを 3 にした場合のサンプリングの不確かさは 26.6%、10.38 $\mu\text{g/g}$ を母平均とするのに対し、母平均が存在すると推定される範囲は 14.5～20.0 ng/g である。しかし、必ずしもサンプルサイズを大きくし、サンプリングの不確かさを小さくしなければならないわけで

はない。仮に、ハム中の亜硝酸ナトリウムの成分規格が上限値として 50 ng/g として設定されていたとするならば、サンプルサイズを 3 としてサンプリングの不確かさを考慮した分析結果に基づき判定しても、規格に適合していることは誤りなく判定することができる。つまり、設定されている規格値等に対して、誤りのない判定をするために必要とされる分析値の信頼性は変わりうるものであり、それを満たすサンプリング計画であれば、性能に問題は無いと考えられる。しかし、本研究でも明らかになったとおり、サンプリング計画の性能に問題がないことを証明するためにも、対象母集団の分布特性をあらかじめ把握しておくことは不可欠である。

・大麦あるいは小麦中のデオキシニバレノール濃度の不確かさの推定

デオキシニバレノール濃度の測定値に含まれるばらつき要因を分離し、併行精度とサンプル間のばらつきを推定した。推定されたサンプル間でのデオキシニバレノール濃度の変動を、1 つの汚染圃場から採取され袋詰めされた大麦あるいは小麦(ロット)中のデオキシニバレノール濃度

分布の母標準偏差として仮定した。また、ロットごとに得られた $n=50$ の試験室代表値の平均値を、当該ロット中のデオキシニバレノール濃度の母平均と仮定した。サンプリングの不確かさ推定の際には、分析精度は一定として不確かさに含めるものとし、サンプルサイズを 3、5、10、16 と変更した場合の影響についても検討した。

まず、中心極限定理により、サンプルサイズの増加に伴いサンプリングの不確かさが小さくなることがわかる。サンプルサイズが 3 の場合のサンプリングの不確かさは、大麦に対する 20.0% が最大であった。その大きさは、昨年度報告したハウレンソウ中での亜硝酸塩濃度のサンプリング不確かさ(14.7%、サンプルサイズは 3)に近い。さらに、母平均が比較的 low 濃度であるため、相対標準偏差とした場合には大きい値として算出されるが、標準偏差とした場合には、最大であった大麦においても、0.6 ng/g がサンプリングの不確かさとなる。同じくサンプルサイズを 3 とした場合に推定される、サンプリングの不確かさを考慮した母平均の存在範囲は、大麦中の母平均が 2.97

µg/g であるのに対し 2.37~3.56 µg/g、トワイズミの母平均が 0.93 µg/g であるのに対し 0.76~1.09 µg/g、ミナミノカオリの母平均が 3.77 µg/g であるのに対し 3.28~4.26 µg/g、農林 61 号の母平均が 1.63 µg/g であるのに対し 1.44~1.83 µg/g と推定される。サンプルサイズを 16 と仮定した場合の母平均の存在範囲は、大麦について 2.67~3.26 µg/g、トワイズミについて 0.81~1.04 µg/g、ミナミノカオリについて 3.45~4.09 µg/g、農林 61 号について 1.51~1.75 µg/g である。サンプルサイズの増加に伴い、母平均が存在すると推定される範囲はより狭くなるが、その幅は小さく、サンプルサイズの増加に伴い増加する分析労力に見合うだけの効果は得られないとも言える。また、この推定は母集団の分布型に正規分布を仮定して行われていることに注意が必要である。デオキシニバレノール等カビ毒のロット内の分布は、その生産菌の分布に依存して局在することが言われており、母集団の分布型は正規分布とならないことも予想される。正規分布を仮定しランダムサンプリングを実行するのであれば、当課題で推定したサンプリングの不確かさ

を指標に、設定される規格値や判定の確からしさを勘案して、より小さいサンプルサイズを決定することもできる。しかし、ロット内に複数の分布が混在する、あるいは母平均を中心とする母分散が左右対称とならないのであれば、その影響はサンプリングの不確かさに現れることになる。従って、母集団の分布型に応じたサンプリング計画を策定しなければ、同じ信頼性を担保し得ない。

E. 結論

生鮮野菜中に残留する各種農薬等のサンプリングに起因する不確かさ

規定したサンプリング計画と手順に基づき、圃場から生鮮野菜を採取し、そこに含まれる各種農薬の濃度を測定した。その結果、個々のサンプル間での測定値の変動は、分析による測定値のばらつき(併行精度)に比べ約 1.5~8.0 倍と非常に大きく、総標準偏差への寄与割合を考えた場合には主たる要因になることが明らかとなった。また、サンプリングの不確かさを推定し、測定値による適合判定の際に考慮することは分析の実効性を損う、現実的ではない行為

であると考えられた。また、本研究ではその推定の際に分析精度を含めたが、何をサンプリングに起因する不確かさであるとするかは十分に議論した上で定義することが必要であると考えられた。

各種生鮮野菜中に含まれる亜硝酸塩濃度のサンプリングに起因する不確かさ

規定したサンプリング計画と手順に基づき、圃場から複数の生鮮野菜を採取し、そこに含まれる硝酸塩の濃度を測定した。その結果、個々のサンプル間での測定値の変動は、分析による測定値のばらつき(併行精度)に比べ大きく、総標準偏差への寄与割合を考えた場合には主たる要因になることが明らかとなった。しかし、Thompson 修正式を用いて一般に推定される測定値の不確かさを指標とすれば、3~10 といった現実的なサンプルサイズにより、十分な精度で母集団中の濃度の平均を推定することが可能であると考えられた。

食肉加工食品に添加される化学物質並びに汚染穀物中に含まれるマイコトキシン濃度のサンプリング

に起因する不確かさ

食肉加工品中の食品添加物濃度のサンプリングの不確かさは、製法の影響を受ける一部の結果を除き、サンプルサイズが 3 の場合でも 7.4~7.6%(RSD)と推定され、正しく製造管理されている加工食品については、極めて微量に残留する農薬等や天然汚染物質また季節変動する食品成分などに比べ、より小さなサンプルサイズであっても、十分な確度で判定可能な分析結果が得られることが示唆された。穀類中のデオキシニバレノール濃度については、サンプルサイズが 3 の場合のサンプリングの不確かさが 11.8~20.0%と推定された。ただしこれは十分に多数のサンプルを系統的にロットから採取し得られた測定値に基づき、正規分布を仮定して推定された結果であり、実際の分布型の推定や、それを考慮したサンプリングの不確かさの推定方法についてさらに検討することが不可欠である。

F. 研究発表

1. 論文発表

なし

2. 学会発表

渡邊敬浩、高附巧、小山陽子、松田りえ子;日本薬学会第129年会(2009年3月26~28日;京都府)

Takahiro Watanabe, Yoshiki Tsukakoshi, Satoshi Takatsuki, Rieko Matsuda;123th AOAC INTERNATIONAL Annual meeting(2009年9月12~17日;米国フィラデルフィア)

渡邊敬浩、塚越芳樹、松田りえ子;第101回日本食品衛生学会学術講演会(2010年9月16~17日;熊本県)

G. 知的所有権の取得状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

不均一分布からのサンプリングのシミュレーションによる最適化

A. 研究目的

複数の圃場や海域で生産、飼育、捕獲された生鮮野菜や海産物等が事前に集荷され、一つのロットやコンサインメントを形成することが、食品の流通に際して一般的に行われる。これらロットやコンサインメントが母集団となる場合には、1つの母集団中に複数の分布が混在することになる。このような状態を不均一分布と呼ぶが、既存の母集団の調査によって、不均一分布であることを解明することは非常に困難である。また、人為的に不均一分布となる母集団を構築することも現実的には不可能である。このような実際に想定されるものの、分析といった物理的手段によって研究することのできない不均一分布を含む特性を持つ母集団を対象とした場合のサンプリング計画の最適化について検討することを目的に、シミュレーション手法を用い、平成20年度:対数正規分布を対象としたサンプリングのモンテカルロシミュレーション
平成21年度:農薬及び硝酸塩濃度、及び不均一分布のサンプリングシミュレーション

シミュレーション

平成22年度:デオキシニバレノール濃度分布を仮定した非ランダムサンプリングのシミュレーションに係る研究を実施した。

B. 研究方法

対数正規分布を対象としたサンプリングのモンテカルロシミュレーション

モンテカルロシミュレーションはCrystal Ball 7(Decisioneering Inc.)により行った。試行回数は全て20,000回とした。

母平均が0.5~1.5、母標準偏差が0.2の正規分布から、Nを3、5、10、25、50としてサンプリングした結果の分布を求めた。判定基準を1とし、サンプル平均が1以下の場合に合格するもととして合格率を求めた。対数正規分布についても同様のシミュレーションを行った。ただし、母標準偏差は0.2及び0.5とした。

上記の分布のロットから採取したサンプルを分析した結果に、分析による変動があることを含めたシミュレーションも実施した。分析による変動の相対標準偏差(RSD)は10%とし、サンプルを個々に分析した場合

と、コンポジットサンプルとした場合のシミュレーションも行った。

農薬及び硝酸塩濃度、及び不均一分布のサンプリングシミュレーション

試行回数を 20,000 回としたモンテカルロシミュレーションを、Crystal Ball 7(Decisioneering Inc.)により行った。

生鮮野菜に残留する農薬及び硝酸塩濃度の分布についてサンプリングシミュレーションを実行するに当たっては、圃場から採取した生鮮野菜 16 試料から得られた分析結果にいくつかの分布を適合し、適合度の高い分布を仮定した。検討する分析対象は、ナス中のピリダリル及びクロルフェナピル、キャベツ中のインドキサカルブ及び硝酸塩、白菜中のエトフェンプロックス及び硝酸塩、大根の葉及び根中の硝酸塩、ホウレンソウ中の硝酸塩とした。分析法は、平成 20 年度及び 21 年度の分担課題「サンプリングに起因する不確かさの推定方法の検討」報告書に示したとおりである。分布の適合は Crystal Ball 7 により行い、アンダーソン-ダーリング検定により評価した。

不均一な分布からのサンプリングシミュレーションでは、母集団分布として正規分布を仮定した。

デオキシニバレノール濃度分布を仮定した非ランダムサンプリングのシミュレーション

分布モデルには、平成 21 年度の分担課題「サンプリングに起因する不確かさの推定方法の検討」で得られたデオキシニバレノール濃度の実測値を用いた。

モンテカルロシミュレーションは Oracle Crystal Ball ; Fusion Edition (Oracle) により行った。

非ランダムサンプリングのシミュレーションのために、以下のモデルを仮定した。

1. 10 のロットから構成されるコンサイメントからのサンプリングを行う。
2. それぞれのロット内でのアナライトの分布は 1 つの正規分布に従っている。
3. 1 個目のサンプルはランダムに採取される(サンプルが採られる確率は、どのロットも 10% である)。
4. 2 個目以降のサンプルが採られるロットの確率は、1 つ前のサンプルが採られたロットに依存して変わる。

1 つ前のロットと同じロットから採られる確率 P_d を用いて、依存の程度を表す。 $P_d=10\%$ の場合は、前のサンプルによらず、全てのロットから等確率でサンプリングされるため、ランダムサンプリングとなる。 $P_d=50\%$ であれば、前のサンプルと同じロットから採取される確率が 50% に上昇し、他のロットから採取される確率は $50/90 = 5.55\cdots\%$ となる。 P_d が大きいほど非ランダム性は高くなる。

5. ロット中の分布のモデルとして、平均 1.5 ppm 、標準偏差 0.3 ppm の正規分布 A と平均 4.0 ppm 標準偏差 0.6 ppm の正規分布 B を仮定した。これらの分布は、大麦中の DON 濃度の分布を想定し仮定した。

6. 10 ロット中に占める分布 B の割合を $0-100\%$ の範囲、 P_d を $10-95\%$ の範囲で変化させた。

7. サンプルは、1, 3, 5, 8, 10, 16 個を採取することを想定し、それぞれの個数の平均値の分布と統計量を出力とした。

8. シミュレーションの回数は 10000 とした。

C.D. 研究結果及び考察

対数正規分布を対象としたサン

リングのモンテカルロシミュレーション

・サンプル平均値の分布

シミュレーションでは、ロット内の分析対象量の分布として 3 種(A, B, C)を使用した。3 種の分布はいずれも、母平均を 1 とした。A は $SD=0.2$ の正規分布で、分布の 95% が $0.6\sim 1.4$ の範囲に含まれる。B は A と同じパラメータの対数正規分布である。母平均は 1 であるが、分布の最も高い点は 0.96 付近で A の正規分布よりもやや小さい。分布の 95% が含まれる区間は $0.68\sim 1.48$ で A よりも大きい側に移動しており、平均値 1 あるいは最大確率値 0.96 に対して、対称ではない。C は対数正規分布で平均 1、 SD は 0.5 である。 $SD=0.5$ では分布の最高点は 0.8 付近で B よりも小さい側に移動し、 95% 区間は $0.37\sim 2.7$ で非常に広がっている。これら 3 種類の分布を持つロットを仮定し、そこからサンプルサイズ N を 3, 5, 10, 25, 50 を無作為に採取したサンプル平均値の分布の統計量を求めた。

正規分布からランダムサンプリングしたサンプル平均値の分布は、平均値が元の分布の平均値 1 と等しく、

母分散/ N の分散を持つ正規分布となる。シミュレーションの結果、サンプル平均値分布の平均値は全ての N の結果において、0.999 あるいは 1.000 で母平均と同じであった。サンプル平均値分布の SD は $N=3, 5, 10, 25, 50$ において、それぞれ 0.115、0.089、0.063、0.040、0.028 となると期待される。シミュレーションの結果は 0.115、0.089、0.063、0.040、0.028 となり、今回実施したモンテカルロシミュレーションにおいて、ランダムサンプリングが実現されていることが確認された。分布の形を示す歪度及び尖度も正規分布を示す 0 及び 3 付近の値となり、サンプル平均値分布において正規分布が保たれていることも、シミュレーションの妥当性を支持する結果となった。

対数正規分布からのサンプリング結果は、SD=0.2 と 0.5 のいずれも、サンプル平均値分布の平均値は 1.0 付近の値となり、母標準偏差 0.2 の対数正規分布からのサンプリングではサンプル平均値分布の SD も正規分布からサンプリングした場合と同じになった。SD=0.5 の場合には、ロットが正規分布であれば、サンプル平均値分布の SD は $N=3, 5, 10, 25,$

50 において 0.289、0.224、0.158、0.100、0.071 と期待されるが、シミュレーションからは 0.287、0.222、0.157、0.100、0.071 が得られ、 N が大きい時にはほぼ一致し、 $N=3\sim 10$ でも 1% 程度の差であった。このようになりに歪んだ分布からサンプリングしたサンプル平均値の分布のパラメータは、正規分布からの結果と大きな差は見られなかった。

一方、 N が小さいサンプリングから得たサンプル平均値分布の歪度は正の大きな値となり、右に裾を引いた分布となっていることが分かる。 $N=3$ でかつ SD=0.2 の対数正規分布からのサンプル平均値は、正規分布に近い分布を示しているが、SD=0.5 の場合には元の分布に似た、右に裾を引いた分布を示している。 N が大きくなると、歪度、尖度共に正規分布の値に近づき、対称分布となる。

・OC 曲線による評価

OC 曲線は、ロットの平均値に対して合格する確率をプロットした曲線で、平均値が基準値よりも十分低ければ合格率は 100%、高ければ 0% となるので、右下がりの曲線となる。基準値が 1 である分析対象を、正規分布及び対数正規分布するロットか

らサンプリングした場合の OC 曲線について検討した。SD=0.2 の正規分布及び対数正規分布からのサンプリング結果はよく似ており、基準値である 1 を中心として対称なシグモイドとなっている。シグモイドの中間部分の傾きは、N が大きいほど急であり、N=50 ではロット平均が基準値を 3% 上回った 1.03 のロットの合格率は 5% であり、基準値を 5% 超えたロットはほとんど合格しない。N=3 で 5% の合格率となるロットの平均値は 1.17 で基準値を 15% 上回っている。このように、サンプルサイズ N を大きくすれば検査の性能は向上する。SD が 0.5 と大きく、非対称性の強い対数正規分布からのサンプリングの OC 曲線は、全体として、SD=0.2 の場合よりも傾きが緩やかであり、N=50 としても 5% 合格ロットの平均値は 1.12、N=3 では 1.4 である。また OC 曲線は対称ではなく、50% 合格するロットの平均値は 1 をやや越えている。このような分布を持つロットがサンプリングの対象であり、N が小さい場合には、ロット平均が基準値を 30% 以上越えたロットであっても 10% の確率で合格する。

・分析精度の影響

分析による誤差の影響を評価するためのシミュレーションを行った。分析による誤差の分布を平均 1、SD=0.1 の正規分布と仮定した。さらに、サンプリングしたアイテムあるいはインクリメントを個別に分析する場合と、全アイテムを混合してコンポジット試料を作成し、これを分析する場合を想定し、2 種類のシミュレーションを行った。分析による変動が、ロットの SD(0.2) の半分であるため、分析の変動を考慮しない場合と、アイテムを個々に分析した場合の結果は、後者が僅かに広がっているだけで、ほとんど変わらない。しかし、コンポジット試料を分析した場合には、N による差が減少し、10 以上ではほとんど差が見られなくなった。正しく混合されたコンポジット試料においてもその濃度の分布の SD は 0.028 となると期待される。しかし、個々のサンプルを分析した場合には分析による変動が繰り返しにより減少するのに対し、1 回しか分析しないコンポジット試料では、分析の変動が 0.1 とサンプル平均値の SD よりも 3 倍以上大きく、全体の変動に支配的な影響を持つためである。

検疫所におけるモニタリング検査

では、検体採取のための開梱数はロットの大きさにより異なり、添加物の検査では2~8、農薬検査では3~32、袋詰め穀物等のアフラトキシン検査では32~210である。また、添加物及び農薬では、サンプルを混合種区分して1検体を作成し分析する。アフラトキシンではロットサイズに応じて1-3の検体が作成される。ロット内の分布が狭く対称性が良い場合には、コンポジットサンプルとすることにより分析の誤差が支配的となるため、多数のサンプルを採取した効果は失われてしまう。このような場合には、少数のサンプルを個別に分析して平均することにより、より効率の高い検査が行える。逆にロット内の分布が広く対称性が悪い場合には、コンポジットとしたことによる影響は少なく、少数のサンプルを個別分析するよりも、多数のサンプルからコンポジットを作成する方が効率の高い検査となる可能性がある。

サンプルサイズN毎に、合格率が5%、10%、20%、80%、90%、95%であるロットの平均値についてみると、Nが10以上では合格率5%~95%の範囲はあまり変化せず、サンプルサ

イズ増加の効果はあまり顕著ではない。さらにコンポジットの分析では、Nを大きくした効果はほとんど見られない。このような分布を持つロットでは少数のサンプルを個別分析した場合に、検査の効率が良くなる。対数正規分布ロットからのサンプリングでは、コンポジットサンプルであっても、Nの増加と共に合格率5%~95%のロット平均の範囲は減少し、サンプルサイズを大きくした効果が現れている。

現実の食品中の分析対象物の分布が明らかになっている例は少ないが、工業的な製品である加工食品中の添加物等は、管理された状態で製造されるので、分布が狭く対称性の良い分布を示すと考えられる。現状では、大きなロットの添加物は8個をとってコンポジットとしているが、ランダムサンプリングが実現されるならば、3個を個別分析する方がより効率の高い検査となる可能性がある。一方、生鮮品中の農薬やアフラトキシンは対称性が悪く、広い分布を示す可能性が高い。このような場合は、なるべく大きなNとすることが望ましいが、多数のサンプルを個別分析することは現実的ではない。しかし、

このような分布から採取した場合には、コンポジットサンプルであっても N の増加に効果が見られるので、なるべく多数を採取することには意義がある。モニタリング検査でアフラトキシンは非常に多数のサンプルを採取してコンポジットとしており、現実的な方法と考えられる。

・サンプリングに起因する不確かさ

今回検討した分布では、母集団の標準偏差とサンプルサイズ N から計算したサンプル平均値は、実測値と概ね一致した。サンプリングに起因する不確かさを、サンプル平均値の分布の SD として考えるならば、元の分布の SD が明らかであれば、対数正規分布のような非対称分布でも、サンプリング不確かさは推定可能である。しかし、真値が存在する確率が 95% である区間という、不確かさ本来の定義とするなら、元の分布の平均値と SD だけでなく、分布の正確な形を知らなくては不確かさは推定できない。ロット内の分布の SD を信頼性高く推定するためには、ロットから採取した多数のサンプルの分析値が必要である。管理下にありロット内標準偏差が一定と考えられる場合には、一度推定した不確かさ

を連続したロットに適用可能であるが、分布が変動すると考えられるロット、あるいは複数のロットを混合したコンサインメントの場合に不確かさを推定することは困難と考えられる。

農薬及び硝酸塩濃度、及び不均一分布のサンプリングシミュレーション

・野菜中の農薬及び硝酸塩分布からのサンプリングシミュレーション

H20 年度の本研究班成果として得られた、実際に圃場から採取した生鮮野菜に残留する農薬濃度をデータとし、シミュレーションを実施した。まず、ナス中のピリダリル及びクロルフェナピル、キャベツ中のインドキサカルブ、白菜中のエトフェンプロックスの 16 個の測定値をヒストグラムにより解析した。その結果、ピリダリル、クロルフェナピル、エトフェンプロックスは、低濃度側に極大があり、高濃度側に裾を引いた非対称の分布を示すが、インドキサカルブは中央に極大があり、他の 3 農薬よりは対称に近い分布を示した。次に、これらのデータに連続分布を適合し、アンダーソン-ダーリング

検定により適合度が良い分布 4 種類を選択した。

また、H21 年度の研究成果として得られた生鮮野菜に含まれる硝酸塩濃度についても同様の解析を行った。大根の葉及び根、ホウレンソウ、白菜、キャベツ中の硝酸塩濃度の測定値をヒストグラムにより解析した。その結果、農薬濃度の分布と比較して、硝酸塩はいずれの作物においても偏りが少なく、対称性の良い分布を示した。さらに、これらのデータに連続分布を適合し、アンダーソン-ダーリング検定により適合度が良い分布 4 種類を選択した。

正規分布ではない母集団であっても、そこから採取した十分多数のサンプル平均値は正規分布することが知られている。そこで、歪みの大きい農薬濃度分布からサンプリングしたサンプル平均値と、対称な硝酸塩濃度分布からサンプリングしたサンプル平均値の分布を比較し、どの程度のサンプルサイズにおいて平均値の分布が対象となるかを検討した。

上記の検討により明らかになった、それぞれの実測値分布に適合した分布から、サンプリングをくり返し、平均値を計算するシミュレーション

を実施した。その結果の一例として、エトフェンプロックスの濃度分布に対数正規分布を適合した場合と、大根葉中の硝酸塩濃度分布に對数正規分布を適合した場合とについて以下述べる。硝酸塩のシミュレーションではサンプル数 3 あるいは 5 で、その平均値はほぼ対称な分布となったが、エトフェンプロックスのシミュレーション結果では 25 サンプルの平均値でも非対称な分布を示した。

歪度は分布の対称性の指標となり、対称分布では 0 となる。大根中の硝酸塩濃度では、仮定した対数正規分布の歪度は 0.518、エトフェンプロックス濃度分布に仮定した対数正規分布の歪度は 1.89 であった。サンプル数を増加させると、どちらの場合もサンプル平均値分布の歪度は低下し、対称性が向上した。しかし、元々歪度の大きいエトフェンプロックスの濃度分布からのサンプリングでは、サンプル数を 25 としても平均値分布の歪度は 0.376 で、硝酸塩の分布においてサンプル数を 3 とした場合と同程度であった。

正規分布からのサンプリングでは、平均値の標準偏差は母集団の標準偏差 / サンプル数の平方根となること

が知られている。エトフェンプロックスについて計算した結果から、非対称な分布からサンプリングしても、その平均値の RSD%は正規分布を仮定した場合にほぼ一致することが明らかとなった。

エトフェンプロックスの平均濃度は 0.03 $\mu\text{g/g}$ であり、Thompson の修正式に従えば、分析の室内精度は 11-14 RSD%と予測される。サンプリングによる平均値の変動を、分析による測定値の変動レベルと同等とするためには、16 個以上の抜き取りが必要となる。この個数であれば、分布の対称性も改善され、正規分布仮定による統計的な推定が適用可能となる。

硝酸塩濃度は野菜により 1000-5000 $\mu\text{g/g}$ の範囲であり、分析の室内精度は 2.2-3.8 RSD%と推定される。エトフェンプロックスと同じように、サンプリングによる平均値の変動を、分析の変動レベルと同等とするためには、16 個以上の抜き取りが必要となると考えられた。

・不均一な分布からのサンプリングシミュレーション

前項では、実測した結果から分布を推定し、そこからサンプリングし

た結果の平均値の分布をシミュレーションで推定した。このシミュレーションでは、ロット内には単一の連続分布が存在しており、そこから完全にランダムに試料を採取したと仮定した。シミュレーションに用いた分布は、単一の圃場から採取した野菜の測定値分布から推定されており、この仮定は現実的である。しかし、実際の検査対象となる生鮮品のロットでは、複数の圃場から採取されたものが混合されている事も多い。本項では、ロット内に複数の分布が存在する場合のサンプリングを考える。

最も簡単な場合として、ロット内に 2 つの正規分布がある場合を想定した。分布 A の平均値は 0.5、分布 B の平均値は 2 とし、標準偏差はいずれの分布も 0.1 とした。分布 B(平均値=2)の全体に占める割合が、10%、20%、30%、40%、50%となる場合の、ロット全体の平均値は 0.65、0.8、0.95、1.1、1.25、標準偏差は 0.46、0.61、0.69、0.74、0.76 である分布からのサンプル平均値の分布を考えた。

上記の 5 種類の分布から、サンプルを 3、5、8、10、16 個抜き取り平均した値の分布をシミュレーションで求めた。その結果、2 つの分布が

重なった不均一ロットからサンプリングした平均値は、複数の極大を持つ分布を示した。例えば、分布 A と B からサンプルを 3 個採取するとき、どちらの分布からいくつのサンプルが採られるかには

A	B
3	0
2	1
1	2
3	0

の 4 種類の場合がある。分布に現れる極大の数はこれに対応しており、それぞれの極大の位置は元の 2 つの分布の平均値から、相対的な大きさは元の二分布の存在割合から計算される。

2 つの分布の割合が等しいロットから 3 個をサンプリングした場合の平均値分布には、4 つのピークが認められ、極大の位置は $0.5(=0.5 \times 3/3)$ 、 $1(=(0.5 \times 2 + 2 \times 1)/3)$ 、 $1.5(=(0.5 \times 1 + 2 \times 2)/3)$ 、 $2(=0.5 \times 3/3)$ である。相対強度は二項分布に従い、低濃度側から 1:3:3:1 の比率となる。分布 B の存在確率が 10% の場合には、相対強度が 729:243: 27:1 となり、B から 3 個がサンプリングされた場合に対応する、極大位置が 2 であるピークはほとんど

認められない。

ランダムサンプリングが行われる限り、このような不均一な分布からのサンプリングでも、サンプルサイズによらずサンプル平均値の分布の平均値は変化しない。また、サンプル平均値の標準偏差も元の分布の標準偏差とサンプルサイズから予測される値と一致している。しかしながら、このような分布から抽出した個々のサンプルにおける標準偏差の範囲は非常に広い。例えば、分布 B の存在確率が 10% の分布の標準偏差は 0.46 である。この大きさの標準偏差を持つ正規分布から、3 サンプルを採取したときのサンプル標準偏差の 95% 区間は 0.073-0.883、16 サンプルでは 0.299-0.534 である。一方、B が 10% の不均一分布からサンプルサイズを 3 として 50 試行した時の、サンプル標準偏差の範囲は 0.024-0.962 であった。サンプルサイズを 16 とした場合には 0.063-0.695 の範囲となり、正規分布よりも大きな範囲となった。また、16 個のサンプルに分布 B からのサンプルが 1 つも含まれない確率は 20% 程度、10 個のサンプルでは 35% 程度である。従って、採取したサンプルをコンポジ

ットとせず個別に分析したとしても、存在確率の小さい分布の存在を見逃す可能性がある。また、サンプル中に存在確率の小さい分布からの試料が含まれていても、連続な非対称分布と解釈される可能性が高く、サンプル測定値から元の分布の不均一性を推定することは難しい。

分布 B が 10%含まれる場合の平均値は 0.65 であり、基準値が 1 ならばこのロットは基準適合である。20%含まれる場合の平均値は 0.8 であるので、やはり適合ロットである。分布 B が 50%であるロットの平均は 1.25 であるので、不適合である。

分布 B が 10%であるロットからランダムに 1 サンプルを採りその値で判定した場合の合格確率は当然 90% である。ところが 3 サンプルの平均値で判定した場合は 85%が合格となり、1 サンプルで検査した場合よりも誤判定の確率が増加した。8 個をサンプリングした場合には 96%の合格率となった。一般の正規分布からのサンプリングでは、サンプル数が多いほど誤判定率は小さくなる。サンプル数を増加させた時に誤判定率が高まるのは、特異な現象である。分布 B が 20%含まれるロットでは、

この現象が顕著であり、8 以下のサンプル数での合格率は、1 サンプルを用いた合格率である 80%以下であった。また、16 サンプルを用いても合格率は 90%で、10%の誤判定率となった。

母平均が 0.8 で変動が 30%程度の正規分布ロットでは、16 サンプルの平均値で判定すれば 99%合格するが、2 つの分布が存在する不均一なロットでは、誤判定の確率も大きくなった。逆に、分布 B が 40%あるいは 50%含まれるロットは不合格であるが、サンプル数が 10 付近で合格率が上昇する、つまり誤判定率が高くなる現象が見られた。

デオキシニバレノール濃度分布を 仮定した非ランダムサンプリング のシミュレーション

・穀類中のデオキシニバレノール (DON)濃度分布

H22 年度分担研究課題「サンプリングに起因する不確かさの推定方法の検討」により得られた穀類中のデオキシニバレノール(DON)濃度のヒストグラム解析を行った。大麦とミナミノカオリではし、ロット全体に DON が高濃度に存在しておりし、特

に大麦ではサンプル間での DON 濃度が大きく変動していることが明らかとなった。また、大麦ではロット全体の DON 濃度がランダムに変動しているわけではなく、1-20 サンプルでは DON 濃度が高く、20-40 では低い傾向が見られた。トワイズミと農林一号の DON 濃度は低く、変動も大麦ほどは大きくなかった。ヒストグラム解析の結果からも、大麦以外の 3 試料では、中心付近に最頻値があり、両端では頻度が少なくなる分布であったが、大麦では全体に頻度が同程度の分布を示した。

DON 濃度を得るための実験では、サンプリング点はロットとした袋の上から下に向かって、らせん状に等間隔に配置されている。DON の濃度に局在があるならば、特定のサンプリング点付近に高濃度が集中する、あるいはサンプリング点に近いサンプルでは濃度が同程度であるという現象がみられると予想された。大麦において、DON 濃度の分布がロット内で一様ではなく、サンプリング点に高濃度のグループと低濃度のグループが存在することが認められたが、他の 3 ロットではこのような明らかな傾向は認められなかった。そこで、

ロット内でのサンプル濃度の関係をより明らかに示すために、自己相関を求めた。自己相関は、1 つの系列とそれを n 個ずらした系列の間の相関関数である。系列を $\{a_1, a_2, a_3, \dots, a_n\}$ としたとき、 $\{a_1, a_2, a_3, \dots, a_{n-1}\}$ と $\{a_2, a_3, a_4, \dots, a_n\}$ の相関は 1 次の自己相関と呼ばれる。自己相関は $-1 \sim 1$ の値をとり、0 次の自己相関は当然 1 である。系列内で隣り合う要素が同じような値になる傾向が強く、離れた要素とは異なる値となる場合には、1 次の相関関数は正の大きな値となる。また、系列内に周期性がありその周期が n であれば、 n 次の自己相関は正の大きな値となる。

自己相関は 1 つの系列をずらして計算されるため、計算に使用される数値の数は、 n の増加と共に減少し、 $n=30$ では要素を 20 個含む 2 系、 $n=40$ では要素数が 10 個の 2 系列の相関となる。従って、相関係数の信頼性は n の増加と共に低下する。

4 種のロット中、ミナミノカオリと農林 1 号から採取したサンプルの自己相関は、1-25 の n の範囲で $-0.2 \sim 0.2$ の範囲の値となった。 n が大きくなると正負の大きな値が見られる

が、一定した傾向は認められない。一方、大麦の自己相関は、 n が 2-5 の範囲で 0.4 程度の相関が見られた後に徐々に低下し、20 付近で -0.6 という強い負の相関となり、再び上昇する傾向を示した。自己相関を解析した全 50 サンプル中、1-20 サンプルでは DON 濃度が高く、20-40 では低い傾向が見られ、長い周期の変動傾向がある。この周期は 40 であるので、その半分の $n=20$ で強い負の相関が現れたと考えられる。系列が 50 よりも長ければ、 $n=40$ で強い正の相関が現れたと期待される。また、これは、2-5 の範囲での 0.4 程度の相関は、隣り合う 5 個程度のサンプルは同じような値となる傾向があることを示している。トワイズミの自己相関は、 $n=15$ 付近に正の相関が現れ、その半分の 8 付近には負の相関が見られる。また、15 の 2 倍の 30 付近にも相関係数のピークが認められる。これは、周期長が 15 の周期性があることを示しており、ロットからのサンプリングが 15 を周期とするらせん状に行われたことから、袋の向かい側にあるサンプル間の濃度差は大きく、同じ側に縦に連なったサンプルの濃度は近いことが推測される。

以上の結果から、特に均一化処理をしないで採取した穀物ロット中のデオキシニバレノール汚染濃度は、周期的に変動する、あるいは高濃度部分と低濃度部分が明らかに区別されるような分布を持つことが実証された。

・ランダムサンプリングと非ランダムサンプリングのシミュレーション

サンプリング計画では、採取するサンプルの個数、あるいはインクリメントの大きさと個数が定められる。サンプリング計画に従って採取されたサンプル中濃度の平均値、あるいは不良率の信頼性は、サンプリング計画によって決まる。これには、ランダムサンプリングが行われているという前提条件が必要である。しかし、ランダムサンプリングでなければ、サンプルから得られる平均値の信頼性が、どのように変化するのかは明らかではない。ロット中でのアイテムの大きさが十分に小さく、アナライトの分布がロット中で均一であれば、ランダムではないサンプリングでも、ランダムサンプリングでも、得られたサンプル平均値は同じような分布を持つと予想される。

H21 年度に実施した本課題研究では、ロット中のアナライト濃度の分布が均一ではない場合に、ランダムサンプリングを行った時のサンプル平均値の分布について検討した。ランダムサンプリングを行っても、少数のサンプルからの平均値は正規分布とは異なる特異な分布を示した。本年度は、不均一な分布からランダムではないサンプリングを行った場合の、サンプル平均値の分布をシミュレーションで検討し、ランダムサンプリングと非ランダムサンプリングから得られるサンプル平均値を比較した。

・非ランダムサンプリングシミュレーションの結果

まず、正規分布 B のみが存在している、つまり均一な分布をしているコンサインメントから、ランダムサンプリングをした場合と、非ランダムサンプリングをした場合の、サンプル平均値の分布を比較した。ランダムサンプリングした場合の、サンプル平均の平均値は、全てのサンプル数においてコンサインメントの平均値である 4.0 ppm と一致した。また、サンプル平均値の標準偏差はランダムサンプリングの理論に従い、

16 サンプルを採取した場合は元のコンサインメントの標準偏差の 1/4 である 0.15 ppm となった。

Pd=50%及び 90%の非ランダムサンプリングを設定して、同様のシミュレーションを行ったところ、サンプル平均の平均値は、全てのサンプル数においてコンサインメントの平均値である 4.0 ppm と一致した。また、サンプル平均値の標準偏差はランダムサンプリングの理論に従い、16 サンプルを採取した場合は元のコンサインメントの標準偏差の 1/4 である 0.15 ppm となった。

以上の結果から、ロット内のアナライトの分布が均一で正規分布に従っているならば、ランダムでないサンプリングをしても、サンプリングによるサンプル平均値の変動は、ランダムサンプリングした場合と変わらないことが、明らかとなった。

次に、分布 B の割合が 50%のコンサインメントを想定し、同様のシミュレーションを行った。その結果、サンプル数が 3 では、3 つのピークが見られる複雑な形の分布を示すが、サンプル数が 8 個では単峰型の分布となった。Pd=50%としたシミュレーションの結果からは、サンプル 8 個