

表 2 穀類等のロットサイズと検疫数量の関係

A) 精米、モルト等

<50kg	20%
50 -2000kg	10kg
2000 -5000kg	15kg
5000 -10000kg	20kg
10000-20000kg	30kg
20000-60000kg	40kg
60000-120000kg	50kg
≥120000kg	60kg

B) 粟、クルミ等

<300kg	10%
300-800kg	30kg
800 -2000kg	45kg
2000 -7000kg	60kg
7000-20000kg	90kg
20000 -10000kg	150kg
100000-500000kg	250kg
500000-2000000kg	400kg
≥2000000kg	600kg

C) いちよう、カシューナッツ等

<60kg	10%
60-1000kg	6kg
1000 -4000kg	8kg
4000 -20000kg	12kg
20000-70000kg	20kg
70000 -300000kg	30kg
≥300000kg	50kg

<60kg	10%
60-1000kg	6kg
1000 -4000kg	8kg
4000 -20000kg	12kg
20000-70000kg	20kg
70000 -300000kg	30kg
300000-900000kg	50kg
≥900000kg	70kg

乾燥果実は、以下による。生鮮食品よりさらに少なく、小さなロットで10%、大きなロットでも最大7kgと決められている。

<5kg	10%
5-500kg	0.5kg
500 -4000kg	1kg
4000 -20000kg	2kg
20000-70000kg	3kg
70000 -160000kg	5kg
≥160000kg	7kg

表 2.1 中華人民共和国向け精米の輸出検査実施要綱

荷口の大きさ(ton)	検査数量
≤0.12	5%以上
0.12-20	6kg 以上
20 -70	10kg 以上
70 -500	15kg 以上
500 -2000	30kg 以上
2000-4000	45kg 以上
4000-10000	60kg 以上
≥10000	80kg 以上

表 2.2 輸入家禽類の採取量

輸入検査申請数量(ton)	検査材料の採取数
≤5	30 以上
5 -20	50
20-40	70
≥40	90

表 2.3 圃場での確認

面積(ha)	サンプルする木の数	合計サンプル数
<2	10	50
2-3	12	60
4	15	75
5-8	20	100
9-16	25	125
17-20	30	150
20 以上	1ha あたり +1	

表 2.4 EU の穀類のサンプリング法

ロットの大きさ(ton)	サブロット数	インクリメントサンプルの数	試験室サンプルの重さ
>1500	500ト ン毎 ton>	≥300	10
300-1500	3	300	10
100-300	100ト ン毎ton	100- 300	10
50-100	5	100	10
20-50	25トン 毎	100	10
10-20	1	60	10
3-10	1	40	6
1-3	1	10	4
0.05-0.5	1	5	2
0.05	1	3	1

各インクリメントは 100g とする

表 2.5 EU の乾燥イチジク以外のサンプリング計画

ロットの大きさ(トン)	サブロット数	インクリメントサンプルの数	集合サンプルの重さ(kg)
≤0.1		10	1
0.1-0.2		15	1.5
0.2-0.5		20	2
0.5-1.0		30	3
1.0-2.0		40	4
2.0-5.0		60	6
5.0-10.0		80	8
10.0-15.0		100	10

15トンを超える場合には、サブロットを15-30トンとして、100のインクリメンタルサンプルを10kg収集する。

図 3 JAS 法における検査方法

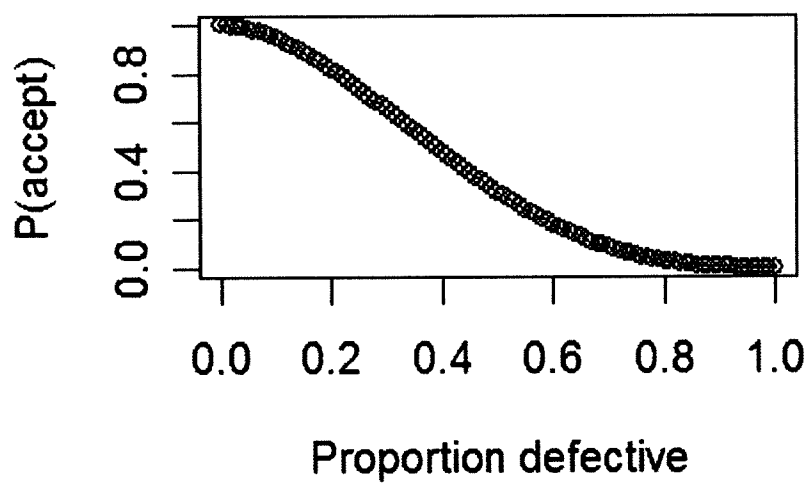


図 3. 1. 1 なみ検査の OC 曲線 飲食料品及び油脂についての検査方法 並検査、小型容器、検査荷口 35000 以下

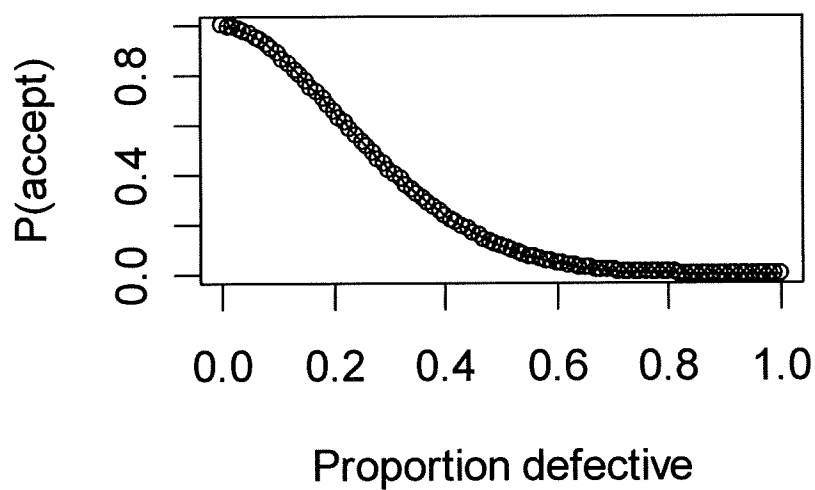


図 3. 1. 2 同、検査荷口、35001-240000

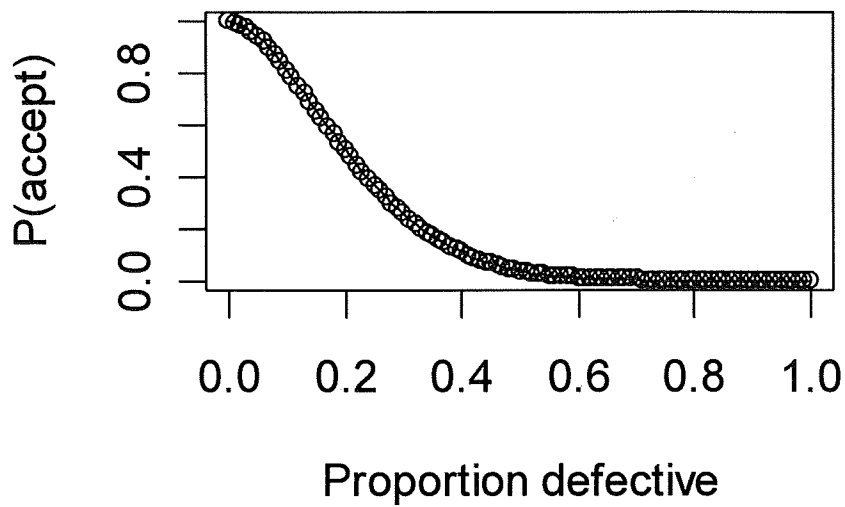


図 3. 1. 3 同 検査荷口 240001以上

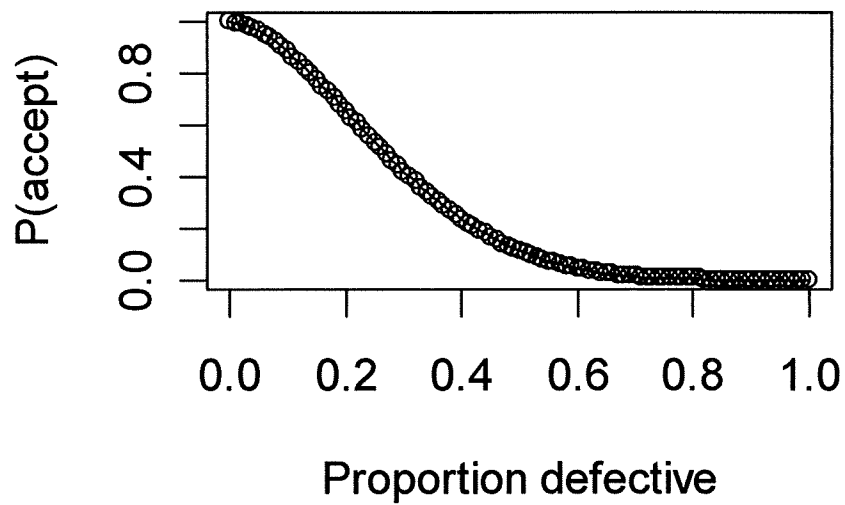


図3. 2. 1 飲食料品及び油脂についての検査方法 きつい検査、小型容器、検査荷口35000以下

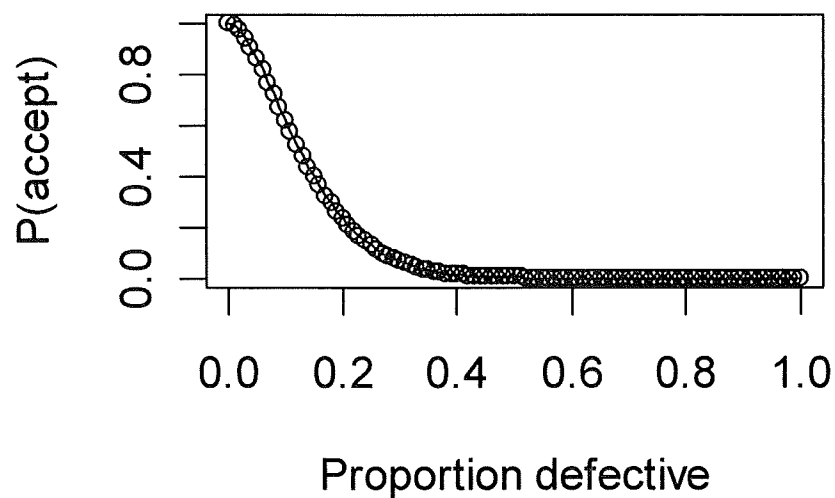


図3. 2. 2 飲食料品及び油脂についての検査方法 きつい検査、小型容器、検査荷口35001以上

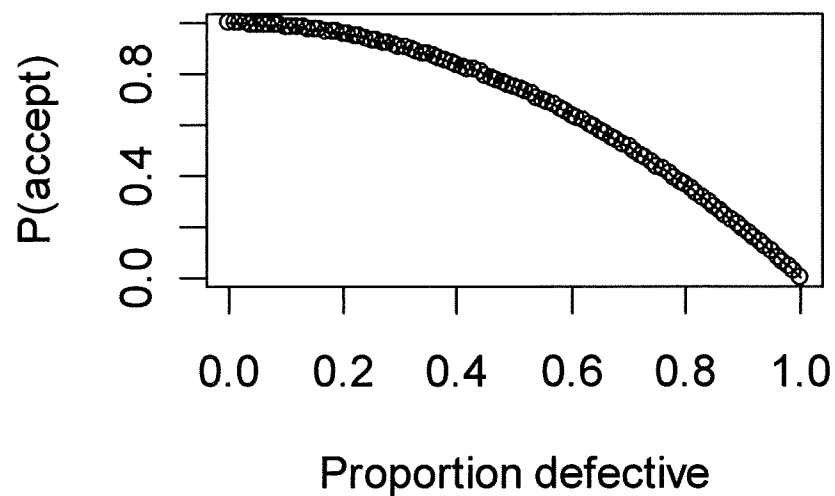


図3. 3. 1 飲食料品及び油脂についての検査方法 ゆるい検査、小型容器、検査荷口35000以下

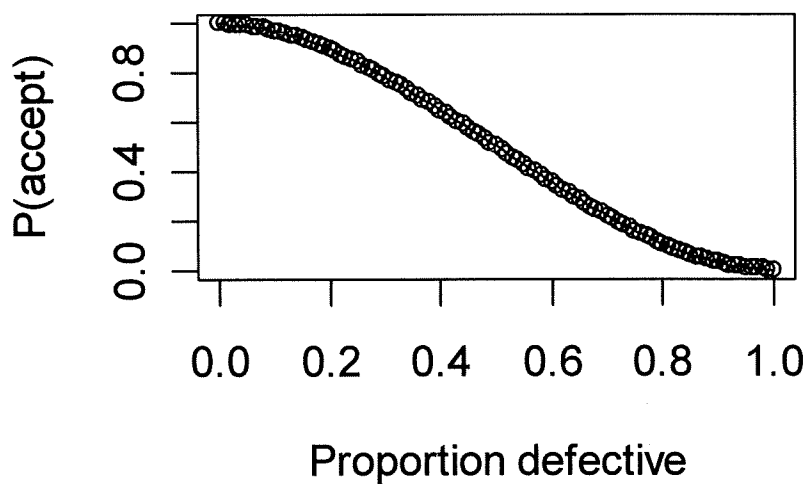


図3. 3. 2 飲食料品及び油脂についての検査方法 ゆるい検査、小型容器、検査荷口35001以上

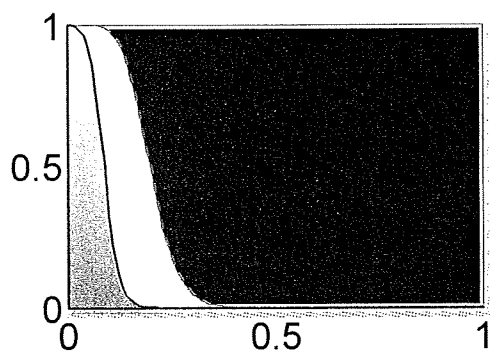


図3. 4. 1 不良率と緩い検査、なみ検査、きつい検査になる確率。赤がきつい検査、緑がなみ検査、青が緩い検査の確率を表す

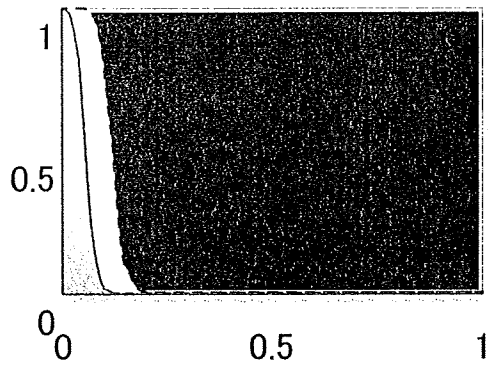


図3. 4. 2 不良率と緩い検査、なみ検査、きつい検査になる確率。赤がきつい検査、緑がなみ検査、青が緩い検査の確率を表す

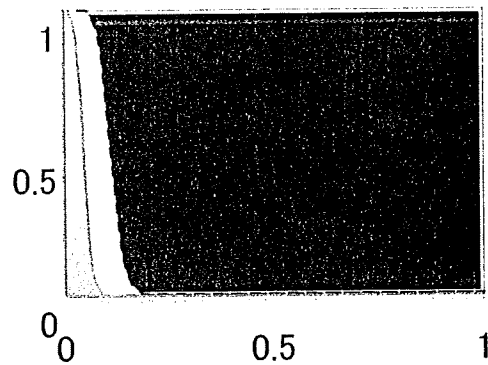


図3. 4. 3 不良率と緩い検査、なみ検査、きつい検査になる確率。赤がきつい検査、緑がなみ検査、青が緩い検査の確率を表す

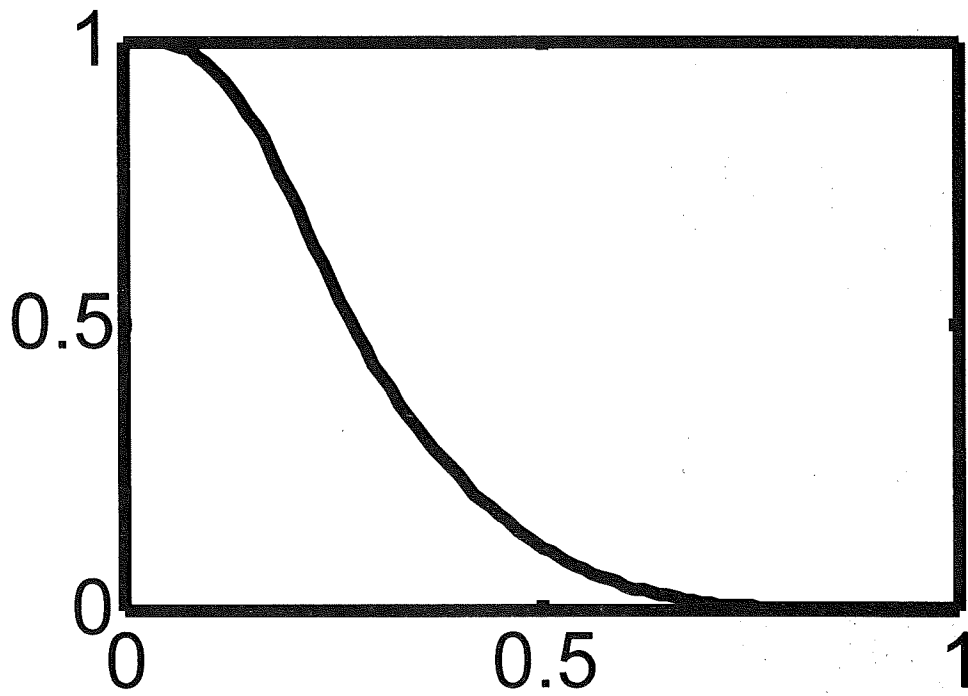


図3. 5. 1 飲食料品及び油脂についての検査方法トータルでのOC曲線 小型容器、検査荷口35000以下

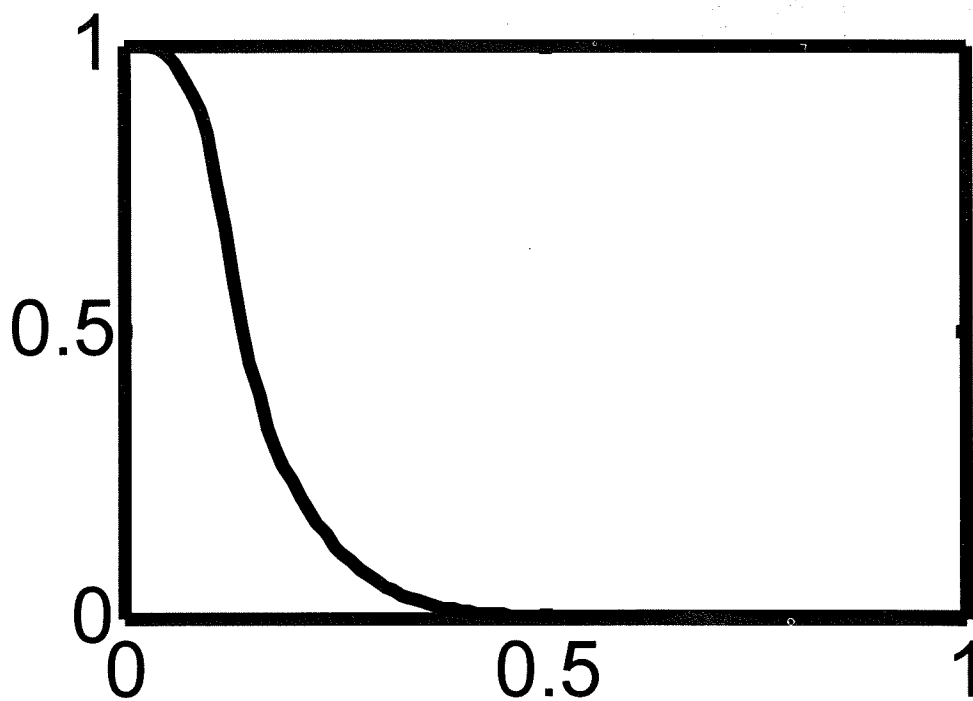


図3. 5. 2 飲食料品及び油脂についての検査方法 小型容器トータルでのOC曲線、検査荷口35001以上240001未満

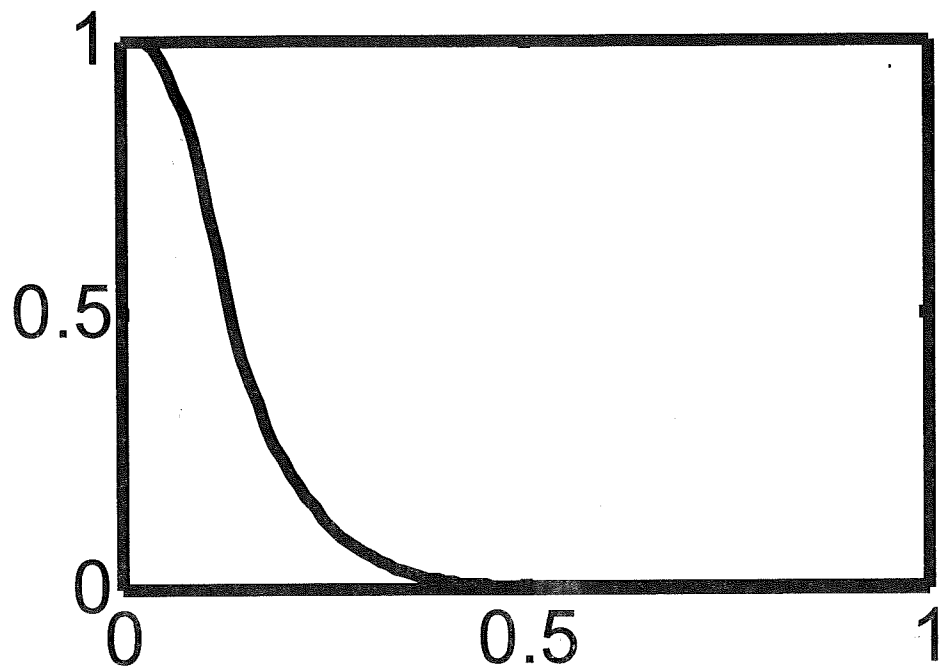


図3. 5. 3 飲食料品及び油脂についての検査方法 小型容器、検査荷口240001以上

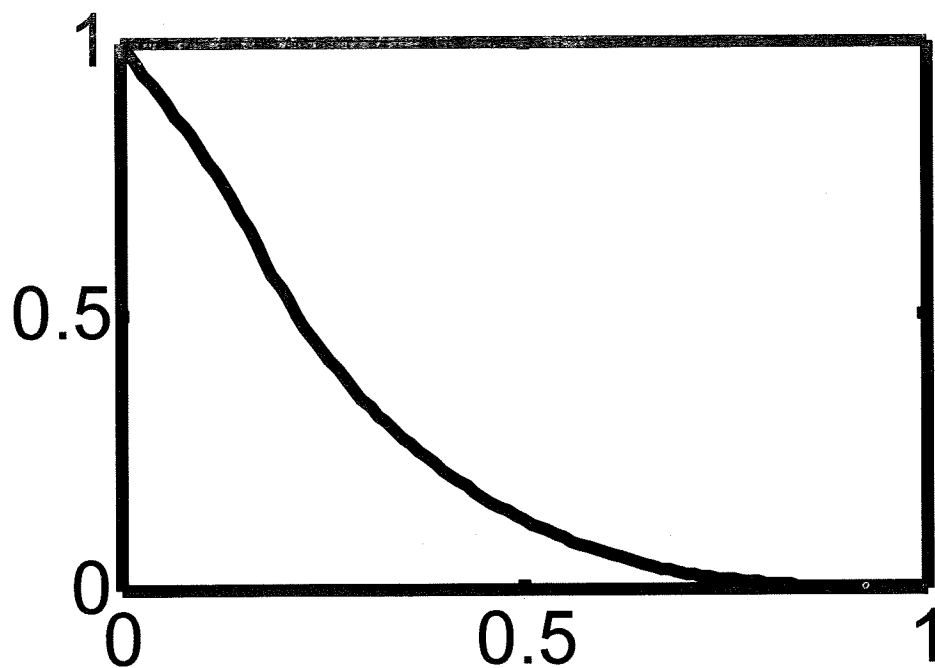


図3. 6. 1 飲食料品及び油脂についての検査方法 大型容器、検査荷口1000以下

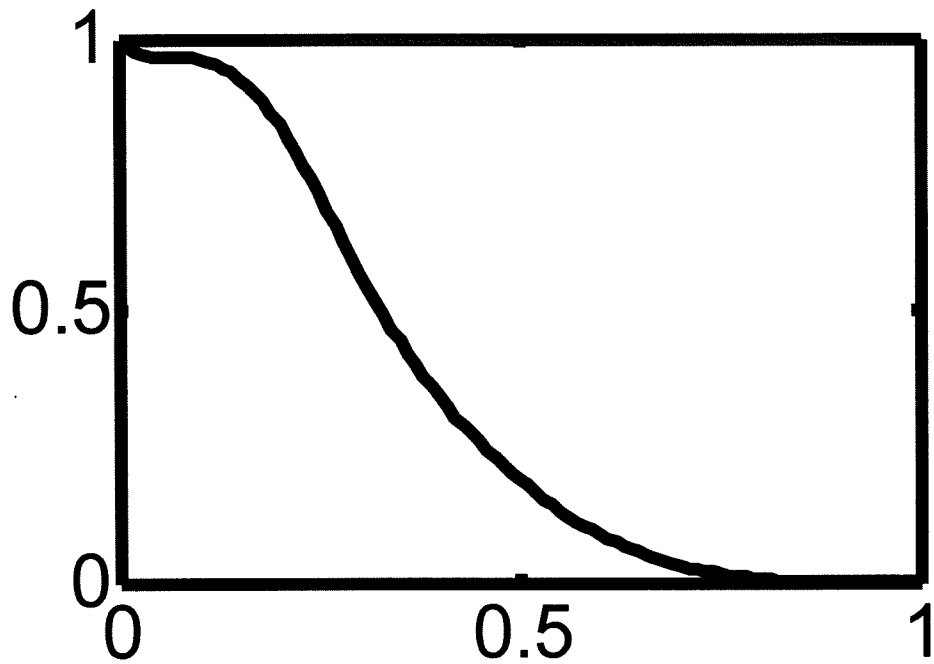


図3. 6. 2 飲食料品及び油脂についての検査方法 大型容器、検査荷口1001以上5000以下

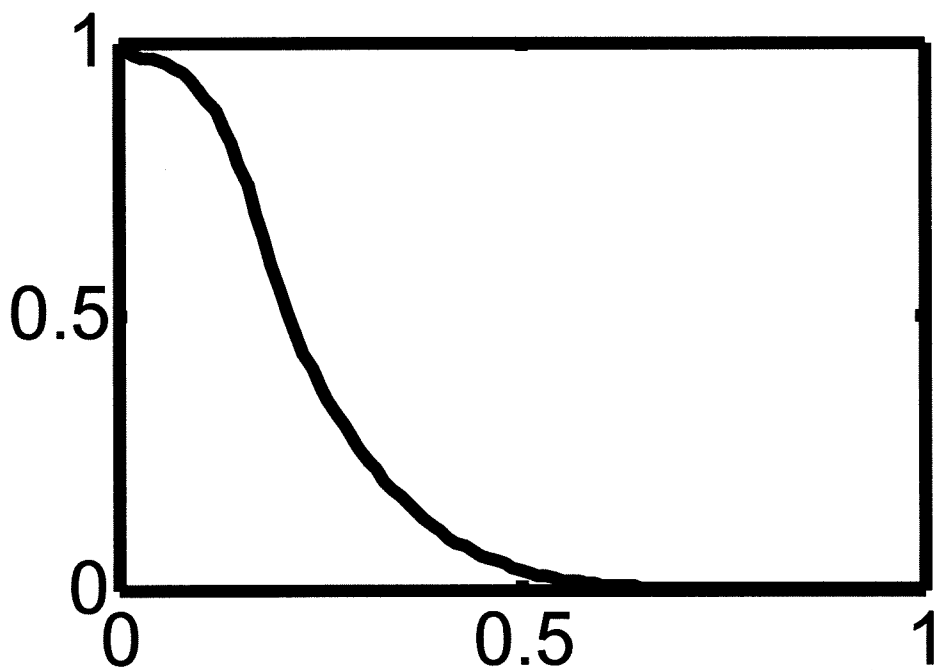


図3. 6. 3 飲食料品及び油脂についての検査方法 大型容器、検査荷口5001以上35000以下

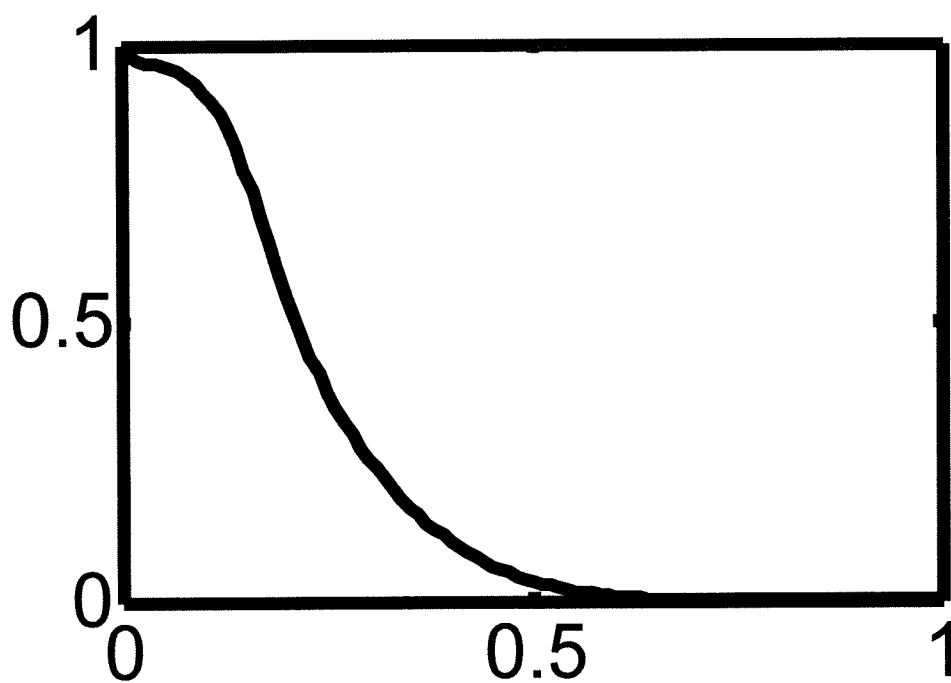


図3. 6. 4 飲食料品及び油脂についての検査方法 大型容器、検査荷口35001以上

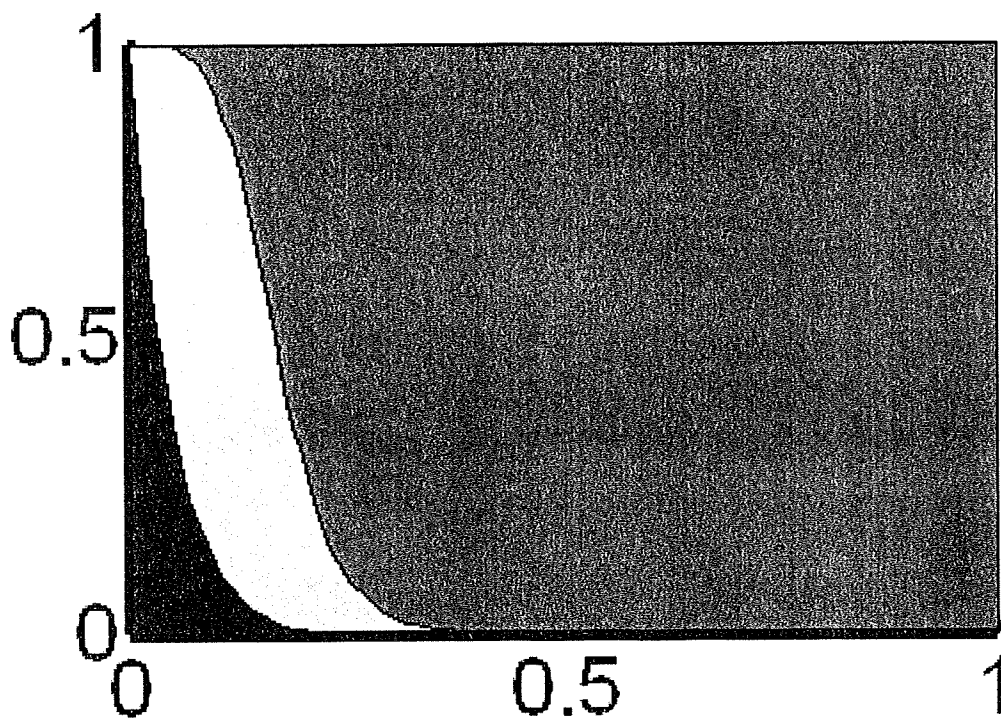


図3. 7, 1 飲食料品及び油脂についての検査方法 の切り替え、黒がゆるい検査、淡灰がなみ検査、濃灰がきつい検査を示す。検査荷口1000以下

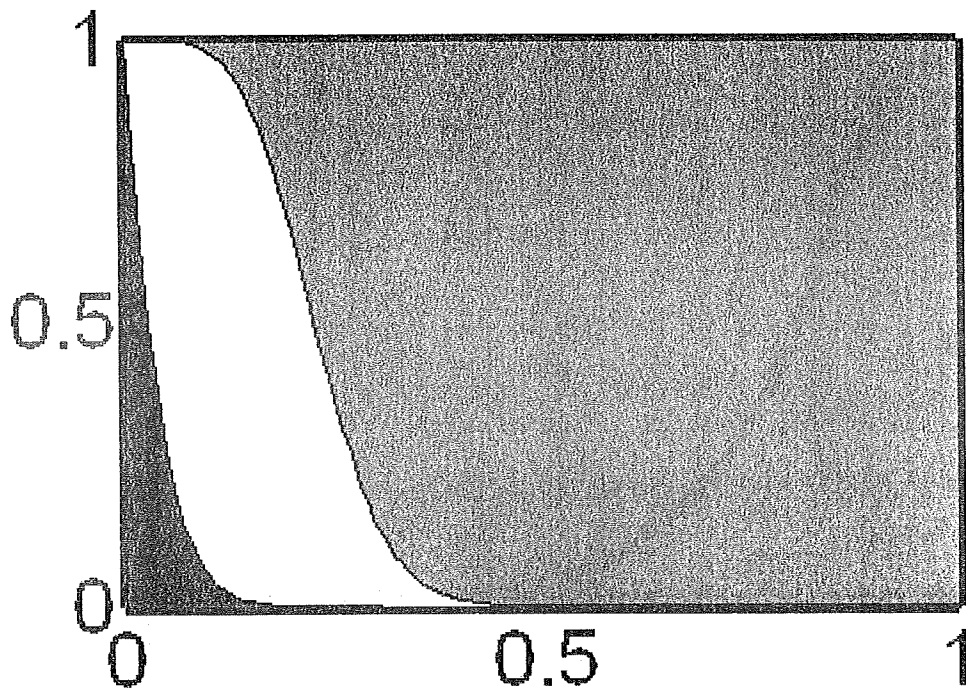


図3. 7. 2 飲食料品及び油脂についての検査方法 の切り替え、検査荷口1000以上5000以下

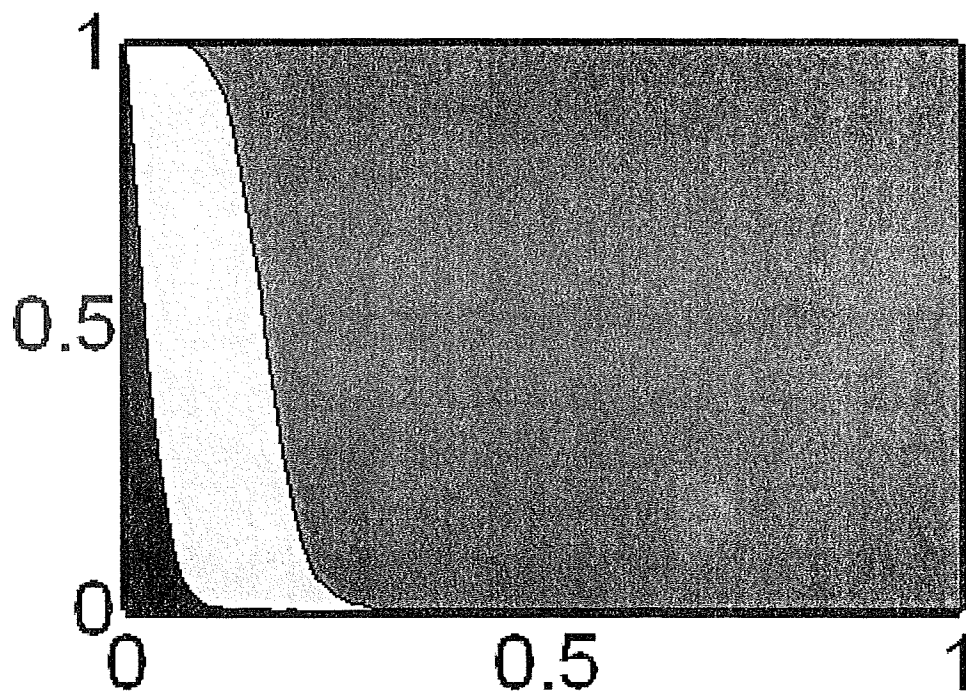


図3. 7. 3 飲食料品及び油脂についての検査方法 の切り替え、検査荷口5001以上

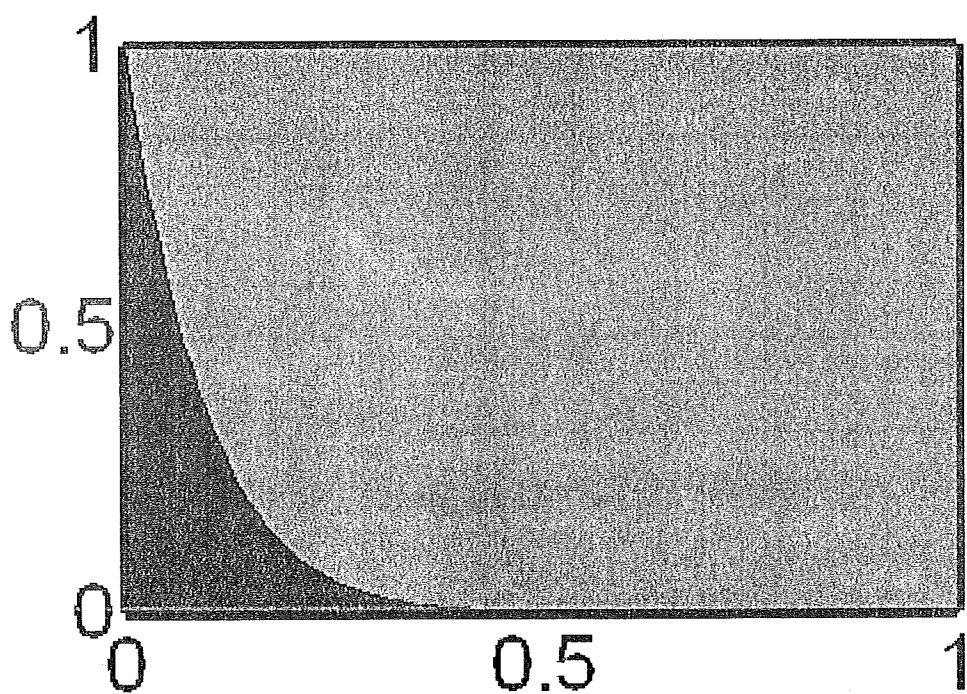


図3.8.1 飲食料品及び油脂についての第1種と第2種検査方法の使用割合（検査の回数）、検査荷口1000以上5000以下

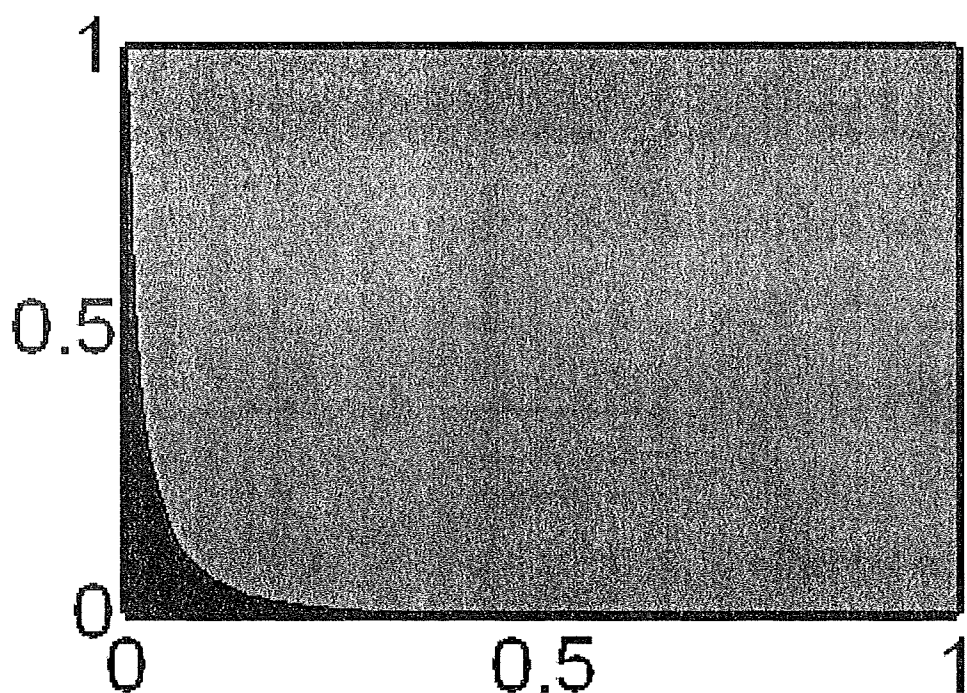


図3.8.2 飲食料品及び油脂についての第1種と第2種検査方法の使用割合（日の長さ換算）、検査荷口1000以上5000以下

表 3.9 熟成ハム、熟成ソーセージ類及び
熟成ベーコン類についての検査方法におけ
る抽出個数（第1種検査方法）

検査荷口 の大きさ	熟成ハム、ベ ーコン	熟成ソーセ ージ
2-1300	1	1
1301-3200	1	2
3201-8000	2	3
8001以上	3	4

表 3.10 JAS法における最小サンプリ
ング量

異性化液糖及び砂糖混合異性 化液糖	500g
植物性蛋白	300g
削り節	50g
ハンバーガーパティ	150g
チルドハンバーグステーキ	80g
調理冷凍食品	150g
醸造酢	150ml
トマト加工品	190g
食用精製加工油脂	900g
豆乳類	180g
マーガリン類	900g
乾めん類	200g
農産物漬け物	100g
チルドミートボール	80g
ブドウ糖	400g
結晶ブドウ糖	200g
ショートニング	900g
精製ラード	100g
煮干し魚類	100g
にんじんジュース	160g

生タイプ即席麺	190g
食料缶詰	50g
パン粉	150g
濃縮果汁	100g
果実飲料	200g
炭酸飲料	1個
ベーコン	1個
即席麺	1000g
マカロニ	50g
ドレッシング	200g
風味調味料	100g
乾燥スープ	200g
ソース	200g
醤油	200g
食用植物油脂	500g

検査におけるサンプリング計画並びに手順の
ハーモナイゼーションに関する研究

II. 分担研究報告書

分担課題 4.

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

松田りえ子

平成 21 年度厚生労働科学研究補助金 食品の安心・安全確保推進研究事業

検査におけるサンプリング計画並びに手順のハーモナイゼーションに関する研究

研究分担報告書

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

研究代表者 渡邊敬浩 国立医薬品食品衛生研究所食品部第三室長
研究分担者 松田りえ子 国立医薬品食品衛生研究所食品部部長

研究要旨

検査対象となる特性値が、理想的な正規分布に従わない分布を持つロットからサンプリングした場合に、サンプルの分析値に基づく検査結果の性質を知るため、生鮮野菜中に残留する農薬濃度及び硝酸塩濃度の分布からサンプリングした、サンプル平均値の分布とサンプリング個数の関係を検討した。農薬濃度は非対称の分布を示し、その相対標準偏差も大きいですが、サンプル数を増加させることによりサンプル平均値分布の対称性は向上し、正規分布に近づいた。サンプリングによる分析結果の変動を分析による変動と同程度とするためには、農薬、硝酸塩共に、16 程度の抜き取りが必要であった。

また、平均値が異なる 2 つの分布が存在しているロットからサンプリングを行った時のシミュレーションを行った。サンプル平均値は多数のピークを持つ複雑な分布を示したが、ランダムサンプリングされている限り、このような場合もサンプル平均値の期待値は母平均に一致した。一般に検査結果によって判定を誤る確率は、サンプリング個数の増加と共に減少するが、複数分布が共存するロットにおいては、検査の結果の誤判定率がサンプリング個数増加に伴い増加する現象が見られた。

研究協力者 渡邊敬浩 国立医薬品食品衛生研究所食品部

A. 研究目的

食品中の有害物質の検査においては、ロットから採取したサンプルの平均値の期待値が、ロットの母平均と一致する、つまりロットを代表するサンプルを採取する事が重要である。母集団がどのような分布であろうとも、ランダムなサンプリングをすればサンプル平均値の期待値は母平均と一致する。しかし、サンプル平均値の分布があま

りに広ければ、期待値が母平均と一致していたとしても、検査結果の正当性は保証されない。サンプル平均値の分布の幅を小さくするためには、サンプルサイズを大きくすることが行われる。

残留農薬検査のように、定量値として結果が得られる検査では、一般にロット中の分析対象物濃度は正規分布していることを仮定して、サンプリングサイズが設定されている。ロット内の分析対

象濃度が正規分布しておりその母標準偏差が既知であれば、サンプルサイズ N と、 N 個のサンプルの平均値の分布の標準偏差の関係、及び消費者危険あるいは生産者危険を推定することが可能である。しかし、現実の生鮮品ロットでは分析対象量の分布が正規ではない場合、あるいは複数ロットからなるコンサインメントのように、2 種以上の分布が重なり合った分布を持つ可能性は高い。

本研究では、このように、検査対象となる特性値が、理想的な正規分布に従わない分布を持つロットからサンプリングした場合に、サンプルの分析値に基づく検査結果がどのような性質を示すかについて検討する。昨年度は、サンプリングのモンテカルロシミュレーション法を確立し、分析対象量が対数正規分布するロットからのサンプルサイズ N を規定したときの、ロット平均と合格率について OC 曲線を用いた比較を行った。非対称性の高い対数正規分布からのサンプル平均値は母平均と一致したが、OC 曲線の傾きは、サンプル数 N を増加させても緩やかであり、正規分布と同程度の効率のサンプリングを達成するためには、非常に大きい N が必要となった。以上の結果から、サンプリング計画のためには、母集団の分布の性質を知ることが重要であることが明らかとなった。

本研究の課題である「サンプリングに起因する不確かさの推定方法の検討」では、実際に圃場で栽培された野菜類に残留する農薬、及びそれらの野菜に天然に含まれる硝酸塩を分析して濃度分

布を求めている。これらの実際の分析対象物の分布からのサンプリングした平均値の分布を検討した。さらに、複数のロットから構成されるコンサインメントのように、複数の分布が重なって存在する分布からのサンプリング結果についても、モンテカルロ法により検討した。

B. 実験

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。

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での母集団の分布は、圃場から採取した野菜 16 試料から得られた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。検討した分析対象は、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、分布への適合はアンダーソン-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。

C. 結果及び考察

C.1 野菜中の農薬及び硝酸塩分布からのサンプリング

リングシミュレーション

農産物中の農薬濃度分布

Fig.1 に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 個の測定値のヒストグラムを示す。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。

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。Table 1 に選択した分布及びパラメータを、Fig.2~Fig5 にそれぞれの分布のヒストグラムを示す。

農産物中の硝酸塩濃度分布

Fig.6 に大根の葉、大根の根、ホウレンソウ、白菜、キャベツ中の硝酸塩の 16 測定値のヒストグラムを示す。Fig.1 に示した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。

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。Table 2 に選択した分布及びパラメータを、Fig.7~Fig11 にそれぞれの分布のヒストグラムを示す。

サンプル数のサンプル平均値分布への影響

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。

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。農薬分布の例として、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を Fig.12 に示す。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。

分布の対称性の指標である歪度の変化を Fig.13 に示した。対称分布では歪度は 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。サンプリング個数を増加させると、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。

Fig.14 には平均値の分布の RSD% の変化を示す。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。図中の点線はエトフェンプロックスについて計算した値を示している。このように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とほぼ一致した。

エトフェンプロックスの平均濃度は 0.03 µg/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 µg/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。

C.2 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、濃度分布を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い。

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。統計量及びパーセンタイル値を Table 3 に示す。分布 B (平均値=2) の全体に占める割合を 10%及び 50%としたときの、サンプル平均値の分布を Fig.15 に示す。

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。図に示した分布に現れる

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。

Table 3 に示すように、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。これらのロットから採取したサンプル平均値が 1 以下になる確率、つまり検査の結果合格する確率を Fig. 16 に示す。

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