

## 医薬品の品質管理・リスクアセスメントへの 統計手法の活用

脇坂 盛雄 (株) ミノファーゲン製薬 顧問

### 1 はじめに

統計／確率と聞いて何を想像されるでしょうか？  
受験の時は偏差値で受験する学校を選びました。この偏差値がどういう意味を持っているかを理解している人はすでに統計をご存知な方です。一回だけ買うなら、宝くじと馬券ではどちらが還付金が大きいかを知っている人は、すでに確率の概念があります。

これまで、統計／確率はどこかで少しは学んでいます。私の経験では、高校の数学で組合せ／確率を学びました。大学では統計を学び「優」（成績の優／良／可／不可）でしたが、ほとんど記憶に残っていません。それは統計を何に活用するかの必要性を認識していなかったからだと思います。学びは学びたいとの強い気持ちが必要になります。何のために学ぶかのその学びの先の目的がしっかりしているかどうかです。

製薬会社に就職して、工場の品質管理部に配属になりました。理学部の化学科で研究室は有機合成でしたので、品質管理部の試験は初めてのことでした。そこで、新人らしく「品質管理」を学ぼうと目的をもって大きな本屋さんに行き、品質管理の本を眺めたら、大学時代に学んだ統計／確率でした。「えっ？品質管理は統計？」と驚いたのを今でも覚えています。

試験をやり出した頃は、高校で学んだ化学が一番役立ちました。統計／確率が必要だと実感したのは、「注射剤の不溶性異物対策」を検討し始めた時でした。まずは、官能検査を本から独学で学び、不溶性異物対策に活用しました。官能検査では人が検査するので、人の教育訓練と認定がとても重要になります。標準見本の作成も官能検査では重要になります。それと人は誤検査（ $\alpha$ と $\beta$ の

誤り）もありますので、その対策も必要になります。検査員の評価はこの $\alpha$ と $\beta$ の誤りで判断します。かつ、全数検査は100%でないため、どうしても残存不良が残ります。そのため、母不良率が高い時はどうすれば良いか等検討しました。

その後、日本科学技術連盟の「QCベーシックコース」を学ぶ機会を会社から与えられました。一週間の講義（月）×6回（6か月）で宿題も多いコースでした。大学と違い、統計／確率の必要性を実感していたので学びにも力が入りました。ここで、統計／確率の概念を身に付けたことが、その後のQC/QAの活動ならびに品質の責任者として判断を求められる時の基礎となりました。その時いただいた優秀賞の時計は今も時を刻んでいます。コースを終了後、工場内でのQC7つ道具の研修の講師として全従業員に伝達しました。

### 2 「ナイチンゲールは統計学者だった！」 丸山健夫著より

ナイチンゲールが統計を活用し、いかに看護が必要かをヴィクトリア女王に訴えました。ナイチンゲールは裕福な家で育ち、病の人を助けることに関心を持ち、当時なかった負傷者を手当てする方法を当時ドイツでやっていたのでそこで学び、その後イギリスの病院でその技術を磨き、ちょうどその時にクリミア戦争が起きました。1855年1月に、平均兵力32,393人中、「感染症」による死者は2,761人、このとき戦争での「負傷」による死者が83人で、「その他」の原因が324人。多くが負傷した後の感染症で亡くなっていました。だれも世話をする人がいなく、ナイチンゲールが人を募って、クリミアに出かけて負傷者の看護（この概念はナイチンゲール

が創出した)を行いました。

病人と病人のベットの間隔が狭いと感染症が増え、一定の距離があると感染症が減ることも統計から導きその後の看護に役立てました。

1855年9月に、平均兵力47,751人中、「感染症」による死者は189人、このとき戦争での「負傷」による死者が276人で、「その他」の原因が20人。看護がどれだけ大切かを訴えるために、データを取る、それを表にする、そして見やすいグラフにすることを政治家やヴィクトリア女王に訴えました。数字だけではわかりにくいのをグラフにすることで見ただけでわかるような工夫を多くしました。ナイチンゲールが統計とグラフを活用しなかったら、女王の理解も得られず、看護の概念はもっと後になっていたでしょう。このように先ずデータをグラフ化する。可視化することが起きている事象を理解する第一歩になります。すぐに検定をする人がいますが、まずはデータを眺める、そしてグラフなど活用し可視化する。そして必要なら最後に検定することになります。製造所では検定まで必要としないケースがほとんどです。

### 3 品質保証業務を行う上で必須になるのが統計／確率の知識

医薬品製造において統計はどのような役割、必要性があるのでしょうか？ 30年、品質管理と品質保証の業務経験から、QAの責任者には以下の知識と経験が必須と実感しています。

- ・剤形ごとの品質の重要なポイントを知る
  - ・固形剤：溶出試験、異物対策
  - ・注射剤：不溶性異物試験、無菌性の保証
- ・剤形ごとの製造法／設備を知る
- ・GMP／公定書（日局・薬添規等）を知る
- ・レギュレーション（外国製造所認定／MF／軽微変更・一変申請）を知る
- ・確率／統計の考え方を（バラツキ & OC 曲線）統計計算ではなく、考え方／活用方法
- ・原料／資材メーカーの品質保証／製造方法を知る
- ・多くの失敗事例を知っておく

## 4 医薬品製造で活用されている統計／確率

- ・統計手法がでてくる場面
  - ・管理図／工程能力指数 ⇒ 工程管理／製品品質照査
  - ・F 検定／T 検定／分散分析 ⇒ 試験方法の変更
  - ・抜取試験（OC 曲線） ⇒ 抜取試験方法設定  
⇒ トラブル発生時（どこからどれだけサンプリング）
  - ・官能検査 ⇒ 人が行う検査の信頼性向上
- ・QC7 つ道具
  - 特性要因図（FISH BONE DIAGRAM）、パレート図、ヒストグラム、（層別）、管理図、グラフ（折れ線／棒／円／帯グラフ、レーダーチャート）、散布図、チェックシート
  - データの可視化を行います。数値データではその姿は見え難いので、QCの7つ道具を使って、データ全体を眺めるようにします。

## 5 統計処理をする前にデータを眺める／グラフにする

研究開発部門からデータ解析結果の報告がありました。

- ・A法とB法でF/T検定を行い有意差ありと報告された。
- ・基本統計量（平均値、標準偏差など）の値から有意差はないのでは？
  - 「両者の平均値の差と、標準偏差からは有意差がないかと思うが？」と尋ねた。
  - ⇒ 「再度、確認して報告する」
  - ⇒ 「インプットミスだった。有意差があるとは言えない結果」だった。

統計手法はエクセルなど解析ツールを使えば簡単に解析結果が出ます。しかし、データを見ることをせずに、解析ソフトに入力して、結果を鵜呑みにしていました。こんなことでは、何をしているのかわかりません。統計を使わないといけないのが、統計に使われています。これなら統計を使ったミスで間違った結果を導いていますので、統計など活用しない方がよいことになります。

品質管理／品質保証を30年経験して、品質保証するにはどうしたら良いか？と尋ねられたら、迷わず、

「3 ゲン (現場, 現物, 現実), 5 ゲン (原理, 原則を追加) を日常で実践する」と返答します。データを森として眺めることと、木一本一本を見る、この両方が必要になります。

## 6 何故、統計はセミナーなどで学んでも実際の場面で活用できないのか

- ・統計の数式を学ぶことが主になっている
- ・数学の数式はわからない
- ・実際の場面での活用が結びつかない
- ・概念よりも手法に目が行っている
- ・計算はパソコン (PC) を使えば結果は出てくる  
A (データ) ⇒ PC (ブラックボックス) ⇒ 結果

統計の学びが概念の理解よりも手法の理解が優先されているように感じています。統計の概念や活用場面についての経験が不足しているためなのかもしれません。そのためには。

- ・データを見る
- ・データが正しいか、偏りないか、データの求め方は
- ・データ全体をグラフなどで可視化
- ・基本統計量の確認  
⇒ 解析ツールを使って計算する
- ・層別 (違う母集団のデータが混在していないか考える)
- ・データがどのようなサンプリング/試験で行われたかの確認

## 7 基本統計量というのを見ていきましょう

基本統計量には主に 2 種類; 代表値と散布度があります。

代表値; 分布全体を一つの数で表しているもの

平均値 (average), 中央値 (median), 最頻値 (mode), 最大値 (maximum), 最小値 (minimum) 範囲 (range) 等

散布度; データのばらつきの大さを示したもの

分散 (variance), 標準偏差 (standard deviation), 変動係数 (coefficient of variation) 等

### 1) 平均 (means)

データの値をすべて足して、足したデータ数で割った値

データ; 90, 98, 98, 102, 103, 104, 105

平均; 100

### 2) 中央値 (median)

データを上から順番に数えて真ん中の値

・データ数 (n) が奇数の場合: 真ん中の値

・n が偶数の場合: 真ん中 2 つの値の平均値

データ; 90, 98, 98, 102, 103, 104, 105

中央値; 102

### 3) 最頻値 (mode)

回数が最も多く現れるデータの値

データ; 90, 98, 98, 102, 103, 104, 105 ⇒

最頻値; 98

最頻値の場合はデータ数が多くないと間違った値を示します。

たとえば, 90, 97, 98, 102, 104, 104, 105 ⇒

最頻値; 104

平均値 100 と中央値 102 は変わっていません。

### 4) 最大値 (maximum) ・最小値 (minimum)

標本集団のうちの最も大きな値と小さな値

データ; 90, 98, 98, 102, 103, 104, 105 ⇒

最大値; 105 最小値; 90

### 5) 範囲 (range)

最大値から最小値を引いた値 ⇒ 15

この基本統計量の平均値, 中央値, 最頻値を知るとヒストグラムの姿もある程度想像することができます。平均 = 中央値 = 最頻値がまったく同じとはどういう場合でしょうか? 富士山のように頂上が山全体の真ん中にあり, 左右対称のような場合になります。

それを正規分布と呼んでいます。医薬品の製造所において, バラツキを考えるとときは, 様々なばらつきを正規分布として仮定しても大きな間違いは生じません。この正規分布の意味をよく理解し, 実際の品質トラブル時にバラツキ, 確率を考慮して検討を進めることになります。

正規分布の状況を表す統計量に散布度がありました。分散は標準偏差を求めるためや統計処理をする時に活用しますが, 実際の現場では扱いません。標準偏差と変動

係数が重要になります。

6) 標準偏差

平均値 (=中央値) からバラツキの大きさを表す指標です。ここで知っておきたいことはどのようにして求めるかの式を理解しておく必要があります。

分散:

$$\sigma^2 = \frac{1}{N} \left\{ (x_1 - \bar{x})^2 + \dots + (x_N - \bar{x})^2 \right\} = \frac{1}{N} \sum (x_i - \bar{x})^2$$

標準偏差:  $\sqrt{\sigma^2}$

この式で覚えておきたいことは、個々の値から平均値を引いた値を二乗してその合計をデータ数  $n$  で割り、その平方根 ( $\sqrt{\quad}$ ) を取った値が標準偏差だということだ。

よって、一つでも飛び離れた値があると標準偏差が大きくなります。データを何も考えずに標準偏差を求めている場合、異常値が含まれていると間違った標準偏差の値を活用してしまうことになりま。データを眺めるとはそのような飛び離れたデータがあるのかどうか、全体のデータのヒストグラムがどうなっているかを知った上で、標準偏差の値を活用します。

7) 変動計数 (coefficient of variation)

相対標準偏差とも呼ぶ

$$CV = \text{標準偏差} / \text{平均値}$$

通常は単位のない数値ですが、% で表す場合もあります。

これがよいのは、平均値の値が異なるデータも 100 を基準として表すことができます。例えば、HPLC/GC の含量の分析バラツキを  $CV = 1.0\%$  とか表します。平均値が 100% であれば、99.0% ~ 101.0% は  $\pm 1\sigma$  だとわかります。

8 PV や安定性試験で 3 ロットの 3 に拘る理由

PV は原則 3 ロット、洗浄バリデーションも原則 3 ロットと、GMP では 3 ロットが基本になっています。なぜ 3 ロットに拘るのでしょう。

データ数

1 個 98%  $\Rightarrow$  98% 1 個だけで他はわからない。

2 個 98%, 104%  $\Rightarrow$  データ数が 2 個あると平均値 (101%) がでる。

3 個 98%, 104%, 98%

$\Rightarrow$  データ数が 3 個あると平均値 (100%) 以外に標準偏差 (2.8%) がでる。

すなわち、データ数が 3 個以上だと統計量としての標準偏差が求まるので、バラツキがわかるのです。ではデータ数が増えると何かメリットがあるのでしょうか?

データ数が 1 個の場合とデータ数が増えた場合の真の値は、データ数が多い方が真の値に近くなります。ではデータ数との関係はどうなのでしょう?

バラツキと  $n$  数は  $\sigma / \sqrt{n}$  の関係があり、 $n$  数が増えるとバラツキは減少していきます。データの 95% 信頼区間 = 平均値  $\pm 1.96\sigma$  となり、母平均がこの範囲に入っている確率が 95% であることを意味しています。 $n$  数を増やすことでこの式の  $\sigma$  を小さくすることができ、母平均の推定範囲が狭くなります。 $n = 1$  に比べ、 $n = 3$  だと  $\sqrt{3} = 1.732$  なので、 $\sigma = 2$  の時は、1.15 まで小さくなります。ですので、受け入れ試験の結果を用いて換算仕込み (受け入れ試験の値で計量する値を変更する) をする場合は、 $n = 1$  で試験を行っていると返って試験結果のバラツキで計量値が異なり、返って品質トラブルを招きかねません。その時は、標準品と検体のサンプル数を  $n = 3$  にして誤差を小さくして、かつ試験成績書との比較を行うなどすることが統計/確率を知った上での品質管理になります。

HPLC の適正条件などに、「 $n = 6$  の注入バラツキが 1.0% 以下の時に試験を開始する」と製造販売承認書に記載されている場合があります。 $n = 5$  にしていると当局から  $n = 6$  にしなさいと一部変更申請時などで修正が入ります。では統計的にみて違いはどの程度でしょうか?

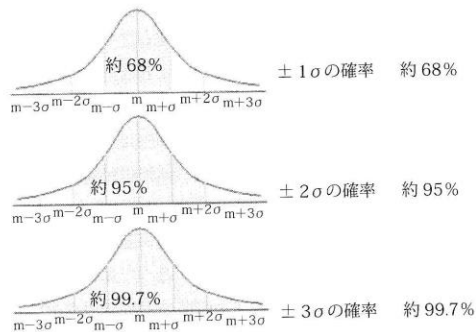
$\sigma = 2.0$  の時、 $2 / \sqrt{5} = 0.894$ ,  $2 / \sqrt{6} = 0.817$  の違いしかありません。当局が拘るのは、分析バリデーションのガイドラインに  $n = 6$  と記載されているからです。実際は大きな影響はありません。このようにバラ

ツキを統計/確率的に物事を眺める、考察できるとその知識が自分のものになり、実際の場面で力を発揮します。

## 9 平均値と標準偏差から、データの発生確率を知る

- 平均値 $\pm\sigma$ の範囲の確率(全体を1) = 0.6827  
 平均値 $+\sigma$ の確率 0.1586 (約15%)  
 平均値 $\pm 2\sigma$ の範囲の確率(全体を1) = 0.9545  
 平均値 $+\sigma$ の確率 0.0278 (約2.5%)  
 平均値 $\pm 3\sigma$ の範囲の確率(全体を1) = 0.9973(1000に3つ)  
 平均値 $+\sigma$ の確率 0.0014 (約0.15%)

正規分布  
<http://www.geisya.or.jp/~mwm48961/statistics/stddiv1.htm>



68%は70%と覚えておいても実際の運用上は問題はありません。99.7%は“1000三つ(せんみつ)”と覚えます。1000個に3つとの意味です。

## 10 事例1) 日局標準品の純度補正がされるようになって

・日局標準品用原料を提供

補正が入るまでは、純度99.3%で調整して提供していました。そして日局標準品として100.0%で販売されていました。つまり、0.7%下駄をばくメリットがありました。標準品管理が国立衛生研究所から当時の公定書協会に移管されました。そこでは標準品に関して、

「99.5%未満の標準品は純度補正係数を設定」のルールが定められました。

その結果、原薬バルクの含量値が0.7%低下することになりました。

そのバルクの日局規格 97.0~101.0%

その結果、原薬バルクの含量値の平均値が97.4%まで下がりました。幸い経年での低下はありません。しかし、分析バラツキの標準偏差は0.4%ほどあります。このことは規格下限まで1σの余裕しかありません。規格の下限を下回る可能性は約15%あります。出荷時の試験では1/6~1/7の確率でOOSが発生します。これは出荷できないリスクがある確率で発生します。かつ、ユーザーで受入試験をして同じ頻度でOOSが発生します。

これは耐えられないと判断し、製造所の合成、試験、業務などの関係者を集め、以下の対応を決めました。

- 1) 純度を高める(蒸留していたので、初留と後留部分をカット)
- 2) 出来高が減り、コストUpになるので純度Upの合成方法を検討する

その結果、ユーザーからの苦情はなく、1年後にコスト削減できる合成方法を確立しました。問題が起きてから動くのではなく、問題が起きる前に問題の発生を予測して早めに動くことができるかどうかです。まさにそれがリスクアセスメントになります。その事象がどういった問題になるかをエンビジョン(envision)することです。それには統計・確率の知識が必須になります。このケースでは規格外がどの程度発生するかをエンビジョンするには統計・確率の知識が役立ちました。±1σ約70%、±2σ約95%、±3σ99.7%(千三つ)の確率であるということを知っているだけで十分です。

## 11 事例2) 溶出試験の製品回収リスク削減には確率の考えが必須

日局には溶出試験の判定方法が2つあります。12個中、10個までが判定値内であればよいとするのと、Q値を用いる方法があります。Q値は新しく設定された判定方法でバラツキに対して緩く、平均値を重視する方法になっています。これまでの経験から、新製品はQ

値設定を行う方が安定性モニタリングで不適合になるリスクが低い傾向があります（設定値によるが）。

バラツキが大きく、1/12の規格外出ている場合は、確率的に溶出試験で7%不適合になります（1/14ロット不適）。

1/12Tの場合；不良率 1/12(p), 11/12(q;1-p)

適合する場合

$${}_{12}C_0p^0q^{12} = 1 \times 1 \times (11/12)^{12} = 0.352$$

$${}_{12}C_1p^1q^{11} = 12 \times (1/12)^1 \times (11/12)^{11} = 0.384$$

$${}_{12}C_2p^2q^{10} = 66 \times (1/12)^2 \times (11/12)^{10} = 0.192$$

合格する確率 = 0.93 (93%)

OOSになり明確なラボエラーが見つからない、製造工程でも問題が見つからない場合。ラボエラーの可能性も否定できないことから、再試験を行います。その時の試験数は  $n = 4 \sim 6$  で行われているところが多いかと思えます。再試験の時は全て適合する必要がありますので、合格する確率は下記になります。

$$n = 4 \text{ で全て合格 } \quad 0.934 = 0.748 \text{ (75\%)}$$

$$n = 6 \text{ で全て合格 } \quad 0.936 = 0.647 \text{ (65\%)}$$

すなわち、1/3 (33%) ~ 1/4 (25%) 不適合になるということです。

統計・確率的に事実を眺めるとは、「時々、追加の6錠/カプセルの試験を行っているということは、いつか溶出試験でOOSが発生して溶出試験不適合になる」を考慮して行動を選択するという事です。そのような製品はいつか安定性モニタリングで溶出試験不適合になり該当ロットの製品回収を行うことになります。

製品回収で当局に説明に行くと、「他のロットは問題ないか?」と尋ねられます。ところが、その根拠データがない場合がほとんどではないでしょうか? そのロットだけ製造で問題があったとか、そのロットの原料に問題があって溶出試験に影響したことを論理的に説明することはハードルが高く、説明できずこれまで出荷した使用期限が残っている全ロットの製品回収になってしまいます。かつ、物流在庫品は問題ないことも説明できず、それも製品回収対象になり即欠品になります。この医薬品に代替品がなく、国民の健康に影響を及ぼす場合は当局も改善されたロットが供給されるまで回収を待つという場合もありますが、代替品がある場合は即製品回収

かつ欠品になります。そこまで先を見通して対策を行うことが統計・確率的に品質問題を眺めることになりま。確率的に物事を考えない品質保証は、運任せの品質保証と言われても仕方ありません。

対策としては下記が挙げられます。

- 1) バラツキを減らす（コーティング方法や打圧調整、添加剤の銘柄変更や粒子径検討など）
- 2) 経年での低下を減らす（処方変更、包装形態変更、保持量追加）
- 3) 判定基準の変更（Q値採用の一部変更申請）

ではどこまで対策すればよいか?

1/15Tの場合；不良率 1/15 (p), 14/15 (q;1-p)

適合する場合

$${}_{12}C_0p^0q^{12} = 1 \times 1 \times (14/15)^{12} = 0.437$$

$${}_{12}C_1p^1q^{11} = 12 \times (1/15)^1 \times (14/15)^{11} = 0.374$$

$${}_{12}C_2p^2q^{10} = 66 \times (1/15)^2 \times (14/15)^{10} = 0.147$$

合格する確率 = 0.96 96%

1/15の場合は4%不適合になる（1/25ロット不適）

⇒溶出試験で毎回1/15の規格外が出ることを統計確率で考察する

$$\Rightarrow \text{OOSになると } n = 4 \text{ で全て合格 } \quad 0.96^4 = 0.849$$

$$\text{OOSになると } n = 6 \text{ で全て合格 } \quad 0.96^6 = 0.783$$

まだ少し不安です。

1/18Tの場合；不良率 1/18 (p), 17/18 (q;1-p)

適合する場合

$${}_{12}C_0p^0q^{12} = 1 \times 1 \times (17/18)^{12} = 0.504$$

$${}_{12}C_1p^1q^{11} = 12 \times (1/18)^1 \times (17/18)^{11} = 0.356$$

$${}_{12}C_2p^2q^{10} = 66 \times (1/18)^2 \times (17/18)^{10} = 0.115$$

合格する確率 = 0.98 98%

1/18の場合は2%不適合になる（1/50ロット不適）

⇒溶出試験で毎回1/18の規格外が出ることを

統計確率で考察する

$$\Rightarrow \text{OOSになると } n = 4 \text{ で全て合格 } \quad 0.98^4 = 0.922$$

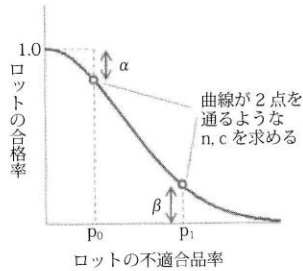
$$\text{OOSになると } n = 6 \text{ で全て合格 } \quad 0.98^6 = 0.886$$

ここまで低下させれば、まず安定性モニタリングで問題になることはないかと思えます。

このように統計/確率を知ることは、判断の根拠になります。

抜き取り試験の基礎となる OC 曲線を理解する

そのロットが合格する割合と抜取数、合格する基準の関係をグラフにしたのが、OC 曲線 (Operating Characteristic Curve) です。



OC 曲線

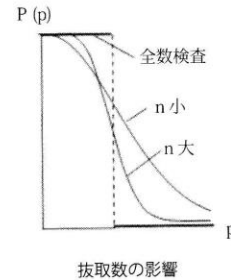
横軸がロットの不良品率で、縦軸がロットの合格率になります。抜取検査では生産者がそのロットを合格させたいロットの不良率の下限を  $P_0$  で表し、消費者がそのロットを不合格とさせたいロットの不良率の上限を  $P_1$  で表します。 $P_0$  を合格品質水準 (acceptable quality level, AQL),  $P_1$  を限界品質 (limiting quality, LQ) と言います。 $n$  はサンプリング数で、 $c$  は判定で許容できる不良品の数です。

このグラフをぜひ覚えていただきたいことは、抜取検査を理解するための第一歩になります。不良率が大きいほどロットが合格する確率は低くなりますが、不良率が大きくても抜取検査では合格することがあること。不良率が小さいほどロットが合格する確率は高くなりますが、不良率が小さくて合格させたいロットが不合格になることがあるということがこのグラフから読み取れます。

不良があるロットを合格させて良いのかと違和感を持たれた方もいらっしゃるかと思います。抜取検査はある程度の不良は認める検査方法であることが前提条件になります。虫・毛髪のように1個でも認めない抜き取り検査をと思う方がいらっしゃいます。1個でも認めないならそれは抜き取り検査でなく、全数検査をしなければなりません。それもその検査方法が100%検出する方法で行うことになります。

$\alpha$  の誤りとは第1種の過誤/生産者危険 (慌て者の誤り: 慌てて合格にすべきものを不合格にする。全数検査機の良い巻き込み率になる) と呼ばれています。一方、 $\beta$  の誤りは第2種の過誤/消費者危険 (ぼんやり者の誤り: ぼんやりして不合格にするべきものを合格させる。全数検査機での良品中に残っている不良品率になる。つまり見逃し率になる) と呼ばれています。 $\alpha = 0.05$  (5%),  $\beta = 0.10$  (10%) と決めています。なぜ、生産者危険が消費者危険より高いのかと疑問に思われるかと思います。なぜなら抜取試験は生産者が設定している抜取基準なので、生産者危険の方を小さくしています。

海外企業と外観や異物の基準を取り決める時、虫・毛髪を日本側が絶対あってはならない、ゼロだと言っても、相手側には通じません。では AQL をいくらにするかとの話になります。AQL は合格させる不良率の上限になるのです。そこで海外とは AQL の一番厳しい、0.01% を選択することで海外企業と合意せざるを得なくなります。



抜取数の影響

このグラフは抜取数  $n$  により、OC 曲線がどう変わるかを示しています。抜取数が少なくなると、グラフの線は寝てきます。すなわち、生産者が合格させたいロットでも不合格になる確率が増え、消費者が不合格にさせたいロットの合格する確率が増えます。検査に要するコスト、 $P_0$  (合格させたい不ロットの良率)、 $P_1$  (不合格にさせたいロットの不良率) をいくつにするかなどを考えながら抜取数を定めます。このように、抜き取り検査は絶対的な評価方法ではなく、確率に左右されている方法であるということです。ですので、全数保証が必要な場合はラインで全数保証の対策が必須になります。異種品混入防止、捺印確認などはラインでの全数保証で対応し

ます。抜取試験はこのように確率の事象に影響されますが、ロットを積み重ねることにより、悪いものは抜取検査で不合格になる確率が高いため、モニタリングとしての機能を持っています。製品品質照査で不適合になる頻度が多いものは、原因究明あるいは品質改善を図ることが重要になります。

## 12 事例3) 匂いの苦情で官能検査が力を発揮

### ・医薬品製造での官能検査

官能検査は医薬品製造所ではあまり重要視されていないのではないかと思います。官能検査は人の五感による検査であり、注射剤の異物検査では特に人の管理が重要になります。ドリンク剤であれば、味/匂い/色が重要な品質特性になります。

測定機器はキャリブレーションが必須です。官能検査は人が測定機械の役割を担っています。人のキャリブレーションとはは何でしょうか？ その人の提供する検査結果が信頼できるかを管理することになります。人の五感を活用する方法なので以下の点が必要になります。

- 1) 検査員の教育訓練と認定
- 2) 標準見本/限度見本の設定と判定基準の明確化
- 3) 検査環境の整備
- 4) 検査方法の標準化

私は入社2年目に注射剤の異物（不溶性異物 & 不溶性微粒子）の試験方法設定を担当しました。日本薬局方製剤総則の注射剤に異物試験が設定されたので、その試験方法で製造所の注射剤の試験ができるかを検討しました。不溶性微粒子試験は顕微鏡法で注射剤をフィルターにろ過してフィルター上の異物を顕微鏡で観察する方法でした。この方法は試験中に異物汚染させなければ、ほとんど人による差はありませんでした。一方、不溶性異物試験は、人が見て異物があるかないかを見る試験方法で、判定はたやすく/明らかに認める異物がないかを人が判断します。そのため、人によるバラツキが大きくなります。

この異物試験方法を見て、

- 1) 人のバラツキはどう管理するか？

2) 観察時間はどうか？

3) 観察環境はどうか？

4) 観察方法はどうか？

が何も明確になっていないことに気がきました。2)～4)は最適な条件や方法を見出して標準化できますが、人のバラツキはどう管理するのだろうか？と思いました。

人による検査=官能検査と思い、手に入る官能検査の本を読み漁りました。

当時の本；

- ・「官能検査入門」佐藤信著
  - ・「工場技術者のための官能検査入門」野呂影勇著
  - ・「官能検査ハンドブック」日科技連官能検査委員会
- 新しいことを始めるにはやはり知識を持つことだと思います。職場には官能検査のことを知っている人がいなかったため本から学ぶことにしました。

たまたま、2日間の官能検査に関する研究発表のセミナーがあったので、上司にお願いして参加させてもらいました。それらにより官能検査で注意すべき点を学ぶことが出来ました。その時の知識と経験はその後で品質保証業務でとても役立ちました。次の事例もその一つです。

クリーム製品の匂いがおかしいとの苦情が1件ありました。そこで、苦情品を製造所に送付して評価してもらいました。評価結果は、保存サンプルも問題ない、苦情品も通常の匂いで問題がないとの報告でした。

匂いの苦情は数件/年間ありました。クリームのため、何度も使っていると他の匂いを付けてしまうことがありました。なので、今回もそのようなものかと思いました。ところが、数日後にまた匂いの苦情が(2件目)来ました。再度、製造所で調査しましたが、同じく問題ないとの回答でした。

おかしいな?と疑問抱きました。それは確率の視点からでした。ところが製造所の回答は「問題なし」でした。年間数件の苦情が一週間以内に2件も来ることは、何か問題が起きている可能性が高い確率になります。もちろん年間数件ある苦情ですが偶然1週間以内に来ることもありえます。

また数日後に苦情が来ました。この時点で、これは何



か問題が起きていると思いました。製造所の匂いの官能検査員の結果がおかしいのではないかと疑いました。そしてメンバーに「官能検査ハンドブック」(必須の本と思購入していました)を渡して、官能検査員を評価する試験方法を考えるように指示しました。一方、製造所から来た3回目の匂い苦情の回答は同様に問題なしとの報告でした。

官能検査員の評価方法を作ってくれました。適切な方法だったので、それで以って製造所の匂いの官能検査員の評価を行いました。これまで「問題なし」と評価した検査員は「問題あり」だったので。微妙な匂いの違いを検出できませんでした。その微妙な匂いの違いを検出できる検査員で官能検査手法により評価したところ、匂いの苦情で来たロットは保存サンプルも同じ匂いの問題を抱えていました。とても微妙なわずかな違いだったので。苦情を出された方は愛用者であり、かつ微妙な違いがわかる優れた官能検査員だったのでした。

保存サンプルをその違いがわかる官能検査員で評価したところ、あるロットからだわかりました。そこで、原料ロット、資材ロット、製造方法の変更/逸脱を調査した結果、クリームの上に乗せるパーチメント紙とキャップを糊付けしている糊に問題がありました。その糊の使用期限管理をしていなかったために、糊が劣化して匂いを出していたことがわかりました。残念ながら糊のロット管理並びに製造記録に糊のロットを記載していませんでした。

クリームは季節商品のために、物流に数十ロットがありました。出荷して良いかどうかの判断が求められます。物流から全てのロットから1カートン取り寄せ、微妙な違いがわかる官能検査員だけで官能検査の評価を行ったところ、不幸中の幸いで在庫としている数十ロットは匂いに問題がありませんでした。その評価結果に基づいて出荷しました。出荷したロットの匂いの苦情は0件でした。まさに、官能検査員の認定を確実に行ったこと、そして評価方法も官能検査の統計手法を活用したことにより、適切に原因分析/評価を早期に実施できました。官能検査の知識を持っていたからできたことだと思います。

官能検査員を評価するのに用いた官能検査の3点識別法

2つの試料を識別するために、どちらか一方を2個、他方を1個、合計3個を検査員に提示して、異なる1個を当てさせる方法です。

組合せの作り方には、次の2つの方法があります。

(1) (A・A・B) 3つの中から同じものを2つ選ぶ

(2) (A・B・B) 3つの中から同じものを2つ選ぶ

～

(6) (A・B・B) 3つの中から同じものを2つ選ぶ

統計・確率を知っているとは、起きている事象を統計・確率の視点から眺めることができるかということです。今多くの製品が溶出試験の安定性モニタリングで製品回収をしています。品質再評価によりほとんどの固型剤に溶出試験が設定されました。試験を作って、季節変動や経年のサンプルを分析したところ、ほとんどの製品で1/12～2/12の判定値外がでました。試験方法を作成した研開の担当者は「日局の試験に合格しているのに問題ありません。何が問題なのですか？」とのことでした。そこで統計/確率の視点から説明し、ようやく理解してくれました。製造方法の変更や添加剤の銘柄変更、打圧変更、包装形態の変更、処方の変更(一部変更申請)などを行い対応を図りました。今溶出試験で製品回収になっている製品はきっと溶出試験を設定して試験したときに1/12～2/12の判定値外がでていたのではないかと推察しています。そのときに対応されていれば製品回収にならなかったではと想像しています。すなわち、統計/確率の知識を駆使して将来起こる問題を想像し、起きる前に対策を行う。これが統計/確率を活用した品質保証になります。