

第5回 評価のスタートのサンプリングの統計/確率を知る (バラツキと OC 曲線)

サンプリングはロット全体の一部からサンプルを取り、それを試験することで全体を推定するために行います。選挙前のアンケートによる選挙結果の予測と同じです。

サンプリングにおいては、計量値と計数値があります。含量や不純物量のように連続した結果が得られるのは計量値と呼び、アンプルの不溶性異物など1本/個と数を数えられるものは計数値と呼び、それぞれの統計/確率の基礎が異なっています。

統計の一週間セミナーを参加され、F/ t 検定や分散分析を習って来られても、肝心の統計的な考え方を学んで来られないことが多いです。現場で統計的な考えができていません。統計/確率は検定の時に使うものと思われているようです。医薬品製造において、難しい統計の考えは不要です。バラツキの概念と OC 曲線 (Operating Characteristic Curve) の概念を身に付けるだけで充分です。統計/確率を数式で考えると数学の知識も必要になり難しいですが、感覚を統計/確率で裏づけることができるだけで充分です。

1. バラツキを知る

1) 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:

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

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です。

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。個々のデータを見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となります。データを眺めるとは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。

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か?

データ数

1個 98% ⇒ 98% 1個だけで他はわからない。

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

3個 98%, 104%, 98%

1.5 0.067

0.061

表の見方は、 $\sigma = 1.5$ であれば、Z が縦軸 1.5 横軸 0 の 0.067 が黒塗りの確率になります。 $\sigma = 1.55$ であれば、Z が縦軸 1.5 横軸 5 の 0.061 が黒塗りの確率になります。日本科学技術連盟の数値表 A にも掲載されています。

また、ネット上 (<https://staff.aist.go.jp/t.ihara/normsdist.html>) にも掲載されています。

バラツキから眺める統計手法（管理図、工程能力指数）

管理図；

管理図は、 3σ 管理です。基本は 3σ 内にあれば、工程は管理されている、 3σ を外れると管理されていない、何か異常が生じていると考える管理手法です。統計/確率を理解していないと 3σ の管理幅から外れた場合は問題で、 3σ の範囲に入っていると問題ないと思っているとそれは統計/確率思考が身に付いていません。管理図 (3σ) 内であっても違うことが起きている可能性があります。逆に管理図外でも問題がない場合 (0.3%) もあるという前提での管理なのです。

工程能力指数；

品質管理の分野において、ある工程の持つ工程能力を管理幅に対して定量的に評価する指標の一つです。

$C_p = \text{規格幅} / 6\sigma$

$C_{pk} = |(\text{規格値上限} - \text{平均値}) / 3\sigma, (\text{平均値} - \text{規格値下限}) / 3\sigma|$ のどちらかの小さい値

規格幅（上限と下限の幅）	工程能力指数	管理状況
4 σ （片側 2 σ ）	0.67	×
6 σ （片側 3 $\sigma = 3\sigma$ 管理）	1.0	△
8 σ （片側 4 σ ）	1.33	○
10 σ （片側 5 σ ）	1.67	◎

C_p は規格幅に対するバラツキの指標であり、 C_{pk} は C_p に平均値の規格幅の中心値からのズレを反映したものになります。どちらも、規格幅に対するバラツキと平均値からの確率を表す指標となっており、値が大きいことは、規格幅に対して工程が安定していることを示しています。工程能力指数と 3σ を関係づけて理解することです。

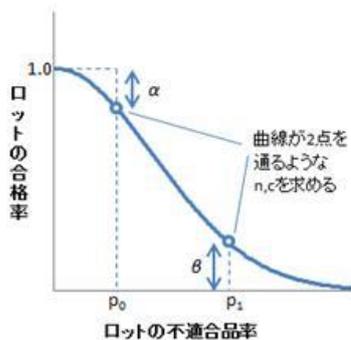
管理図を使うか工程能力指数を使うか。工程の管理には管理図管理、製品品質照査や OOT の設定は規格値との関係が重要ですので工程能力指数の C_{pk} で考察します。

バラツキを知るとは、その事象が起きる確率を予測することができます。OOT を 3σ 管理で行っているが、1/10~20 回に 1 回 OOT がでるので困っているとの話を聞くことがあります。これは統計/確率的な考え方ではその現象はおかしいことになります。なぜなら 3σ 外になるのは 0.3% なのです。330 回に 1 回起きる確率なのです。なぜこのようなことが

起きているか？ それは 3σ の値が不適切だからです。たまたま良かった/季節変動が入っていなかったなどの理由で 3σ の値が小さかったからです。その場合 OOT を工程能力指数で考えると、 $Cpk > 2.0$ だったりします。それであれば、 3σ でなく、 4σ または 5σ で管理すれば良いのです。確かに工程では何か 3σ からの違いが出ているのかもしれませんが、その変動は規格幅から考えたら微々たるものになります。重箱の隅を爪楊枝でつつくようなものです。意味のないことにエネルギーを費やしても品質保証には何のメリットもありません。逆に OOT が出ているのに、毎回「しばらく様子を見る」とのコメントを付記している方が GMP 上問題になります。即ち原因究明を行っていないことになります。仮に原因究明をして改善してもそのメリットは微々たるものです。それにエネルギーを費やすなら他にやるべきことがあるはず。統計/確率の概念を知っていれば、全体を俯瞰し何が重要かを把握できるようになります。

2. 抜き取り検査の意味と限界

OC 曲線/抜き取り検査の理解



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。抜き取り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。

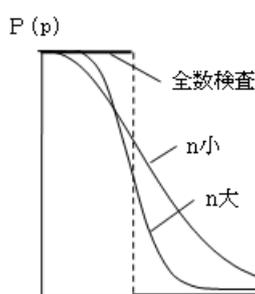
OC 曲線

このグラフをぜひ覚えていただくことが、抜き取り検査を理解するための第一歩になります。不良品率の大きいほどロットが合格する確率は低くなりますが、不良品率が大きくても抜き取り検査では合格することがあること。不良品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、不良品率が小さくても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。抜き取り検査にはこのような性質を持っています。

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。抜き取り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。かつその検査方法が 100% 検出する方法で行うことになります。多くの検査は 100% 検出する方法ではありません。検査したから良品との思い込みは捨てなければなりません。

α 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。全数検査機の良品巻き込み率になる)と呼ばれています。一方、 β 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えます。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。



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 、 P_1 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。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必須になります。

JISZ9015 抜取検査

原料の抜取試験/サンプリングはドラム数に関係なく、1カ所だけのサンプリングで行われている製造所が多いのではないのでしょうか？ 1カ所だけであれば、それは1カ所だけで問題がないとの根拠データが必要になります。下記にその根拠を示します。

- ・ロット内が均質である(どこから採ってもロットを代表している)
 - ・製造方法で均質化工程があり、均質になっているとのバリデーショndataがある
 - ・不均一化を来す作業がない(掻き落とし作業など)
 - ・実際に数多くサンプリングして均質との結果を得る(検証する)
- ・製造所を査察し、コンタミや異種ラベルなどの問題がないことを確認する。
 - ・洗浄バリデーション確認
 - ・高薬理活性を兼用ラインで製造していない
 - ・ドーピング薬剤を兼用していない。兼用している場合は ppb レベルで確認する。原薬製造ラインでドーピング薬剤がコンタミ(GMPの基準以下)したためそのわずかにコンタミした原薬を使った製剤の製品回収が起きている。
- ・ラベル発行管理にミスが生じないことの確認

統計的に根拠を持った抜き取り検査を行うなら、グローバルな基準である JISZ9015 に基づいて抜取数を定めることになります。

統計/確率は難しいと拒否されずに、ぜひ取り組んでいただきたいです。バラツキと OC 曲線を理解することは医薬品製造において必須だと 30 年 QC/QA に携わってきて実感しています。いろいろな問題が起きた時、統計/確率で考えないと品質トラブルの判断ミスを招きます。他社さんになりますが、ファームテック・ジャパンに 7 回に渡って統計/確率を楽しくわかりやすく掲載させていただきました。ぜひ見ていただき、統計/確率を自分の強みの一つにさせていただきたいです。

サンプリングは評価の基本です。サンプリングを何か所から行うか。その時のサンプリング量をいくらにするか、サンプリング方法をどうするか。計量値であればそのサンプリングしたものを粉碎して均質にしてそこから試験に用いる量をサンプリングするかなど、その行為一つひとつに統計/確率の意味があります。サンプリングは評価の出発点であり、それを考えるための手段が統計/確率です。

データインテグリティ (DI) にエネルギーをかけて取り組まれている製造所が増えていきます。ところがそういう製造所の中に、サンプリングを現場に任せている製造所もあります。現場に任すと必ず現場のフィルターがかかっています。良いサンプルだけがサンプリングされている場合もあります。偏ったサンプルを完璧な試験で行っても意味があるのでしょうか？ 改ざん防止/隠蔽防止のために GMP は製造と評価を分けています。ところが現場でサンプリングするとその GMP の大前提が崩れています。日本の GMP では“原則“という言葉で現場でのサンプリングも認められていますが、それを利用し GMP の基本を疎かにしている行為だと言っても過言ではないと思います。DI の前に QC がサンプリングを！