

2つの比率の差の同等性試験に関する研究

2007MI143 水野裕太

指導教員：松田眞一

1 はじめに

現在の後発医薬品の生物学的同等性試験ガイドライン [1] には、試験製剤と標準製剤の生物学的同等性判定パラメータの対数値の平均値の差の 90%信頼区間が、 $\log(0.8) \sim \log(1.25)$ の範囲にあるとき、試験製剤と標準製剤は生物学的に同等と判定するとある。その他の同等性試験ガイドラインも同様で、平均値の差を用いる方法のみ記述されている。

そこで私は、2つの比率の差の信頼区間を用いて同等性を判定する場合にはどのようにすればよいのかという疑問を持った。ガイドラインには比率の差を用いて行う方法については全く触れていないので、比率の差を用いて同等性を判定する方法自体は解明されているのか調べてみた。その結果、比率の差の信頼区間で同等性を判別する方法は、正確にはないということが分かった。比率の差の同等性試験についてあまり研究されない理由は何なのか、そしてそれを行うにはどのような問題点を改善しなければいけないのかということを追及し、比率の差を用いた同等性試験を行うことができるのかどうかを研究することを目的とする。

2 比率の差の問題点

丹後 [2] に詳しく示されているように、比率の差の検定では、平均値の差と異なり信頼区間との対応関係は正確にはない。標本数 n_A と n_B の値が大きく異ならず、成功率 p_A と p_B もほぼ同じ値を示す時、近似的に信頼区間と対応関係があるといえる。しかし、今回の研究では、比率の差の信頼区間を用いて無条件で同等性を判定する方法を考えたいので、この問題を解決する必要がある。

比率の差の同等性試験に関する様々な文献を読んでいくと、ほとんどの文献が信頼区間を作る際に正規近似を用いている。正規近似を用いてしまうと、信頼区間の上限値、または下限値にマイナスの値が出てきたり、1を超える値が出てきたりしてしまう時がある。本来、比率の差の信頼区間を構成するのであれば、 $p_A - p_B$ の値は、必ず $0 \sim 1$ となるので、その信頼区間は、 $0 \sim 1$ の範囲内で作られるべきである。この問題は、同等性試験に限らず、比率の差の信頼区間の構成に関する問題点として、一般的に指摘されている。

最後に残る問題が、比率の差の信頼区間を用いた場合、どのような範囲した時、同等と判定することが出来るのかということである。平均値の差を用いた場合、試験製剤と標準製剤の生物学的同等性判定パラメータの対数値の平均値の差の 90%信頼区間が、 $\log(0.8) \sim \log(1.25)$ の範囲にあるとき同等と判定すると示されていたが、比率の差でもこれは使えるのか考えてみると、ある問題が生じることに気付く。まず、成功率が著しく大きい値となる時、 $\log(0.8) \sim \log(1.25)$ の範囲が大きくなり、上限値にお

いては、1を超えることもある。それに対して、成功率が著しく小さい時は、 $\log(0.8) \sim \log(1.25)$ の範囲が狭くなってしまふ。また、この範囲に入れば同等であると言えるような根拠はない。すなわち、どのような範囲にあるとき同等とするのかということも考える必要がある。

3 Fisher の正確確率検定に対応する比率の差の信頼区間

3.1 導入

先ほど示した問題点を解決する為に、松田 [3] の方法を考えていく。この方法は、オッズ比の信頼区間から比率の差の信頼区間を求めるものであり、Fisher の正確確率検定に対応する比率の差の信頼区間を求めることが出来る。また、この方法は正規近似を用いておらず、正確な比率の差の信頼区間を求めることが出来る。すなわち、この方法を用いることで前節で挙げた 3 つの問題点の内、2 つが解決できる。

ここで、記号を導入し、 2×2 分割表を次の記号で示す。

	有効	無効	計	真の有効率
薬 A	x	z	n	p
薬 B	y	w	m	q
計	k	l	N	

また、 $x \sim B(n, p), y \sim B(m, q)$ とそれぞれ二項分布に従っている二項分布モデルとする。

3.2 比率の差の信頼区間の構成法

オッズ比の信頼区間から比率の差の信頼区間を求める為に 2 つの方法を松田 [3] から以下に引用する。

1. 周辺度数を真値であるかのように扱い、比率に関する次の条件式を設定し、それを基にオッズ比の信頼区間から変換する方法。

$$np + mq = k$$

(この式は「2つの群を合わせた場合の真の比率は k/N である」ということを定式化したものである。)

2. オッズ比に対応する比率の組 (p, q) の中で「2項分布モデルにおいて周辺度数が得られる確率」が最大となるものを求める方法でオッズ比の信頼区間から変換する。

前者の方法は、信頼区間を求める方法が簡便であるが、安易な方法であるという印象を受けてしまう。それに対し、後者の方法は最尤法に似た考え方であり受け入れやすいが求め方が複雑となる。しかし、この二つの方法から求められる信頼区間は一致することが、松田 [3] で証明

されているので、求め方は前者の方法を、意味は受け入れやすい後者の方法をと考えることが出来るので、これらの方法から比率の差の信頼区間を求めていく。

4 同等判定の基準

後発医薬品の生物学的同等性試験ガイドラインに示されている判定基準は、比率の差を用いた場合、適切ではないと考えられるので、新たに2通りの判定基準を考える。1つ目は、判定区間を $[-0.1, 0.1]$ と固定して、この範囲内に入れば同等であると判定する方法であり、2つ目は、判定区間を $[-\sqrt{pq}/5, \sqrt{pq}/5]$ とし、この範囲内に入れば同等であると判定する方法である。これらの判定基準で同等と判定できるのか、次節で検証していく。

5 シミュレーション

2つの薬剤が「同等である」といえる 2×2 分割表のデータを作り、その基準でどのくらい当てはまるのかシミュレーションを行っていく。また、松田 [3] の方法から求める信頼区間の信頼率は 95% とする。

	有効	無効	計	真の成功率
薬 A	x	$n - x$	n	p
薬 B	y	$n - y$	n	p

2つの薬剤が「同等である」といえる 2×2 分割表のデータは、上記の表のように2つの薬剤の真の成功率を共に p とし、A群、B群それぞれの例数を共に n として作る。また、 x は $B(n, p)$, y は $B(n, p)$ とそれぞれ二項分布に従う乱数を発生させる。例数 n は、 $n = 100, 200, \dots, 700$ の7通り用意する。それぞれの例数で100個のデータそれぞれから求める95%信頼区間が何回判定区間に入るのかカウントし、それを1000回繰り返してその平均値を出し、その結果を基にして考察していく。

5.1 判定区間を $[-0.1, 0.1]$ で固定する場合

判定区間を $[-0.1, 0.1]$ としてシミュレーションを行った結果が以下の表である。

各群の例数	$p=0.5$	$p=0.4$	$p=0.3$	$p=0.2$
100	0	0	0	0.061
200	0	1.475	10.558	35.171
300	34.733	35.512	48.208	70.439
400	58.344	61.514	71.519	87.611
500	75.914	77.221	85.044	94.854
600	85.912	87.417	92.337	98.09
700	91.628	93.046	96.175	99.332

表1 判定区間 $[-0.1, 0.1]$ の結果 (単位:%)

このシミュレーション結果から、それぞれの例数が600を超えれば、 $p = 0.5, 0.4, 0.3, 0.2$ の全ての成功率で8割以上のデータが判定範囲に入っているため、比率の差の信頼区間が $[-0.1, 0.1]$ の範囲の中にある時同等と判定できるといえる。このシミュレーションを行うまでは、判

定区間を $[-0.1, 0.1]$ で固定してしまうと、成功率の違いで、判定区間への入りやすさが大きく異なってしまおうと予想していた。しかし、結果から成功率が低くなればなるほど判定区間に入りやすくなることは予想通りだったが、判定区間への入りやすさはそこまで大きく変わらないことが分かったので、成功率が低い時は判定区間に入りやすい方が良いと考えたい場合は、この方法を用いると良いと考えられる。

5.2 判定区間を $[-\sqrt{pq}/5, \sqrt{pq}/5]$ と与える場合

判定区間を $[-\sqrt{pq}/5, \sqrt{pq}/5]$ としてシミュレーションを行った結果が以下の表である。

各群の例数	$p=0.5$	$p=0.4$	$p=0.3$	$p=0.2$
100	0	0	0	0
200	0	0.12	0.525	1.638
300	34.806	32.051	32.437	32.348
400	58.425	58.538	58.038	58.197
500	75.914	74.72	75.026	75.09
600	85.996	85.182	85.43	85.532
700	91.756	91.731	91.85	91.791

表2 判定区間 $[-\sqrt{pq}/5, \sqrt{pq}/5]$ の結果 (単位:%)

このシミュレーション結果から、先に挙げた方法と同様、2つの群のそれぞれの例数が600を超えているとき、 $p = 0.5, 0.4, 0.3, 0.2$ の全ての成功率で8割以上のデータが判定範囲に入っているため、求めた比率の差の信頼区間が判定区間 $[-\sqrt{pq}/5, \sqrt{pq}/5]$ に入った場合、同等であると判断しても良いといえる。この区間は二項分布の分散を利用していることから、判定区間内への入りやすさはどの成功率でも、ほとんど一緒となる。成功率の違いによる判定区間内への入りやすさに違いがない方が良い場合は、この方法を用いるのが良いと考えられる。

6 おわりに

シミュレーションの結果から、各群の例数をかなり多くしなければいけないことが分かったので、比率の差の信頼区間の信頼率を90%としたときについても考えてみた。その結果、2つの群のそれぞれの例数が500を超えているとき、8割以上のデータが判定範囲に入ることが分かった。しかし、信頼率を90%としてもかなり多くの例数を必要とするので、もっと例数を少なくしても判定できるような方法を考えるのが、今後の課題である。

参考文献

- [1] 後発医薬品の生物学的同等性試験ガイドライン:
http://www.nihs.go.jp/drug/be-guide/GL061124_BE.pdf
- [2] 丹後俊郎：『新版 医学への統計学』。朝倉書店，東京，1993。
- [3] 松田真一：『比率の差の信頼区間に関する考察』。計算機統計学，第18巻・第2号，pp95-105, 2005。