

解説【仮説検定とベイズ統計】

- 「統計的に有意」は仮説検定で用いられている用語である。放射線リスク研究では、 $\beta=0$ (リスク上昇なし)を仮説にとり $\beta>0$ (リスク上昇あり)を対立仮説にとって仮説検定をおこなうのが一般的である。(あなたが過剰相対リスク係数を知っているならそれを β と考えるとよい。) 検定の目的は仮説 $\beta=0$ がデータから否定されるかどうかを知ることである。そのために $\beta=0$ の母集団から抽出した標本の分布を考える。この標本のひとつがデータである。標本の分布は β がマイナス無限大からプラス無限大まで広がっている。したがってデータの β がどんなに大きくてもこの母集団の仮説を否定することはできない。そこである程度おおきな値 c を決めておいて、 $\beta>c$ であれば仮説を棄却することにする。「棄却する」という言い方をするのであって「否定する」とは言わない。仮説が棄却されたときに用いるのが「対立仮説が統計的に有意であった」という表現である。 c は標本の分布上で $\beta>c$ の確率が α になるように決める。 α を有意水準という。標本の分布上でデータの β の値より大きい区間の確率を p 値という。したがってデータの β と母集団の β のへだたりが大きくなるほど p 値は小さくなる。放射線リスク研究では有意水準として 0.05 がよく使われており、 $p<0.05$ で仮説を棄却している。仮説が棄却されても仮説の正しい確率が 5%あることを承知しておく必要がある。

信頼区間も標本の分布をもとに考える。信頼区間の目的は母集団の β が確実に含まれる範囲を知ることであるが、これも標本の分布の性質のためにデータから母集団の β の確実な範囲を決めることはできない。そこで母集団の β がある確率で存在する範囲を決めることにする。これが信頼区間である。たとえば 95%信頼区間は標本の 95%で母集団の β が存在する範囲を正しく与えるが、残る 5%の標本では正しい区間を与えない。信頼区間と仮説検定は密接な関係にあり、信頼区間をもとに仮説検定の結論を知ることができる。仮説が棄却されるのは母集団の $\beta=0$ が信頼区間の下限の下に位置するときである。このとき対立仮説 $\beta>0$ が統計的に有意になる。なお、95%信頼区間の下限以下の区間の確率は 2.5%であるので、この仮説検定の有意水準は 0.025 としたものになる。

- 仮説検定は標本の分布をもとにした分析法であるが、これと違って直接、 β の値

を推定する統計学がある。ベイズ統計学である。これは最初から β の分布を考え一貫して β の分布を扱う。したがって $\beta > 0$ (リスク上昇)の確率を直接求めることができる。原爆被爆者の固形がん死亡率の分析で 100mSv 以下でのリスク上昇が統計的に有意でないという報告があるが、ベイズ統計学を使えば、同じデータを使ってリスク上昇 ($\beta > 0$) の確率が 95.3%(概算)であるという結果をえることができる。リスク上昇なしの確率はわずかに 4.7%である。100mSv 以下でもリスク上昇がはっきり見えていることがわかる。仮説検定でリスク上昇が統計的に有意になるのはベイズ統計学で $\beta > 0$ の確率が 97.5%以上のときにあたる。リスクについて仮説検定がいかに鈍感であるかがわかるであろう。

- 仮説検定は米国統計学会が声明した通りで研究の世界に大きな弊害をもたらしているが、ここでは、とくに放射線リスク研究との関係で次の点を強調しておく。第一に、仮説検定はデータのもつ情報のおおくをむざむざ捨てているという欠陥である。天気予報の例で説明すると、ベイズ統計なら「翌日の降水確率は 20%」という予測が可能なのに、仮説検定では「あす雨が降る」という命題を棄却できるかどうか問題にしない。第二に、仮説検定は仮説をどう選ぶかで結論の印象が大きく変わるという欠陥である。100 ミリシーベル以下で固形がん死亡リスクの増加は棄却されないことが国民に宣伝されているが、閾値なし線形モデル LNT を仮説にして仮説検定すればこれも棄却されない。このように仮説検定はリスクの知覚に鈍感であり、仮説の選び方によって自説に有利に誘導できることを示している。ベイズ統計であれば、どちらの仮説が妥当であるかを確率という数値によって客観的に示すことができる。

- 仮説検定とベイズ統計学について詳しくはそれぞれの教科書を参照してほしい。