

# 統計モデルを通じて何を知るべきか

——ロジスティック回帰モデルを事例として——

日本大学 常松淳

Email: vyb03233@nifty.com

## 1. 目的

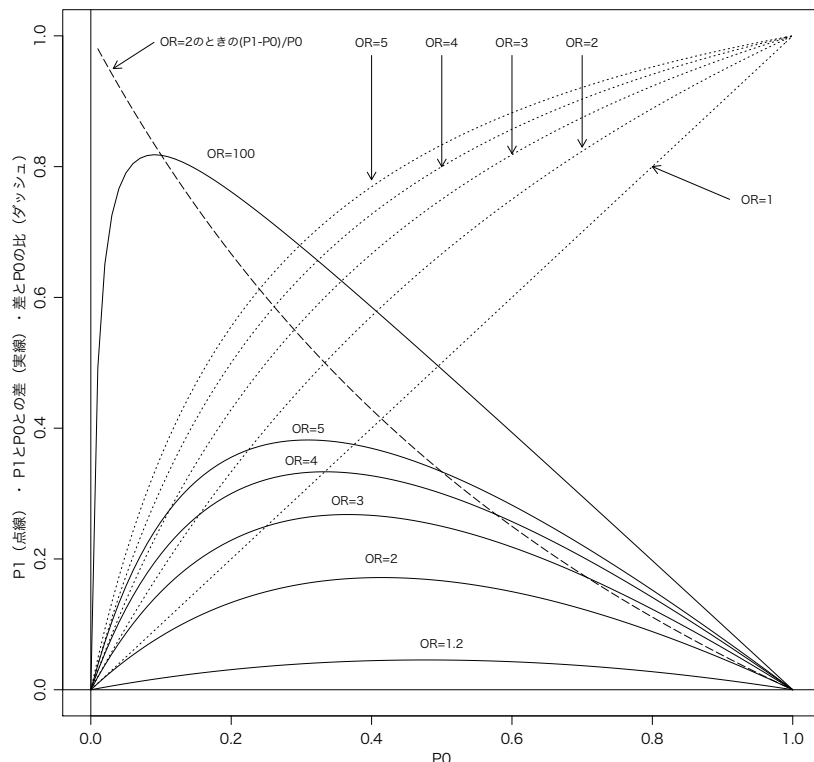
カテゴリカルな変数を分析する手法として二項ロジスティック回帰モデルは社会科学において広く用いられている。回帰系モデルによる分析では、帰無仮説検定を行って有意となった説明変数について個別の“係数解釈”をするというのが通例の研究慣行である。本報告では、ロジスティック回帰モデルの係数解釈にともなう困難を通じて、計量的研究は統計モデルを通じて何を知るべきかという問題について再考する。

## 2. 方法

ロジスティック回帰モデルでは線型モデルと同様の係数解釈ができない。そこでしばしば、ある説明変数を+1だけ変化させたとき、 $\exp(\beta)$ がオッズ比になるという便利な事実が利用される。しかしオッズ比は確率と異なりストレートな解釈が難しいだけでなく、その確率との関係も単純でない。そこで、オッズ比が一定の値であるとき確率はどのように変化しうるのかを数理的に再確認してみる。

## 3. 結果

図は、もとになる確率 ( $P_0$ ) と、それに対して一定のオッズ比にある確率 ( $P_1$ ) との関係を表している。横軸は  $P_0$ 、点線は各オッズ比における  $P_1$ 、実線は確率の差 ( $P_1 - P_0$ )、破線は差の  $P_0$  に対する比を示す。 $P_1$ 、差、比はいずれも  $P_0$  の関数であり、あるオッズ比にある  $P_1$  が  $P_0$  に対してどの程度変化しているかは、 $P_0$  のレベルに応じて大きく異なっている。言い換えれば、オッズ比は推定されたモデルについてきわめて限られた情報しか伝えていないのである。モデル全体の意味を十全に評価するには、説明変数すべてを考慮してモデルの予測値 (確率予測) を見るしかない。



図：P0とP1との関係 [ORはオッズ比]

## 4. 結論

特にロジスティック回帰モデルでは「個々の係数解釈」というかたちに拘泥せず予測値を評価すべきだという点は、諸テキストでも強調されてきた [Long (1997) 等]。それでもなお (統計的に有意となった) 説明変数の係数解釈にこだわることの背景には、研究慣行に加えて、そもそも統計モデルを通じてわれわれは何を知るべきなのかに関する見方の (たとえば、inference か prediction かという) 違いがあるだろう。

【文献】 Long, J. Scott, 1997, *Regression Models for Categorical and Limited Dependent Variables*, SAGE.