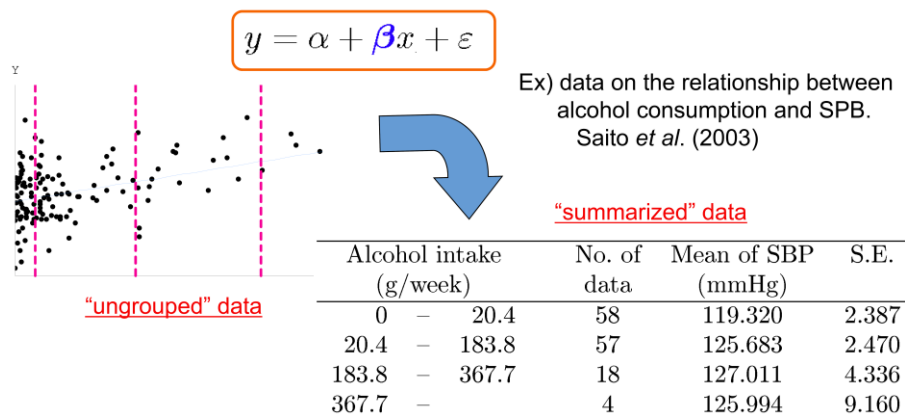


## 区間データに基づくメタアナリシスの方法論と実践

医学・疫学分野の研究において、観察・観測されたデータから種々のリスクを適切に推定し評価することは生物統計学の重要な役割のひとつです。例えばアルコールの摂取と乳がんの発生の関連を調べる場合、アルコール摂取量を曝露量として、その量に応じて乳がんの発生率がどのように変化するか統計的な検討が行われることとなります。しかしそのような研究において、本来連続量であるべき個人々の摂取量などが範囲として調査されたり正確に記録されなかったりすることで、例えばアルコール摂取量として0~20.4, 20.4~183.8, 183.8~367.7, 367.7~ (g/week) のように曝露量が区間データとなることがしばしばあります。それとともに他のデータも各区間で要約されたものが与えられ、解析には少なくなった情報のみが利用されることとなります。このような場合、各群をカテゴリーとした解析が行われるのが一般的ですが、この区間の決め方が各研究でまちまちであると、そこから導き出される結果や数値の解釈も研究間で違ってしまいます。さらに複数の研究結果を集め、それを統合するメタアナリシスを行う際には、この区間データの定義が異なることが結果の併合を技術的に困難なものにします。それにも関わらず、これらの問題をきちんと評価しないまま複数の研究を統合したメタアナリシス研究が行われ、それがエビデンスとして評価されているものも存在しています。

そこで Takahashi and Tango[1]では曝露量にべき正規分布を仮定したうえで最尤推定量によって各区間の代表点を推定することを提案し、線形回帰モデルの下で検討を行い、従来の中点を用いるものよりも精度よく回帰係数の推定を行うことが出来ることを示しました。また非線形関係にある用量反応曲線の推定[2]、カットオフ値が異なるバイオマーカーの値と予後の関連の評価[3]などの問題に対して、同様の区間データに基づく適切な推定方法やメタアナリシスの実践の検討なども行っています。



◎ 各区間の曝露量の代表値を尤度に基づき 7.0, 71.8, 255.0, 530.6と定めることで、  
回帰直線の傾きを精度良く推定できる。

文献：

[1] Takahashi K, Tango T. Assignment of grouped exposure levels for trend estimation in a regression analysis of summarized data. *Statistics in Medicine* 2010; 29(25):2605-2616.

- [2] Takahashi K, Nakao H, Hattori S. Cubic spline regression of J-shaped dose-response curves with likelihood-based assignments of grouped exposure levels. *Journal of Biometrics & Biostatistics* 2013; 4:181.
- [3] Sadashima E, Hattori S, Takahashi K. Meta-analysis of prognostic studies for a biomarker with a study-specific cut-off value. *Research Synthesis Methods* 2016; 7: 402-419.