応用問題 1

(1)

最尤推定を用いて各都市ごとの回帰直線を求めると、次のような結果が得られた。

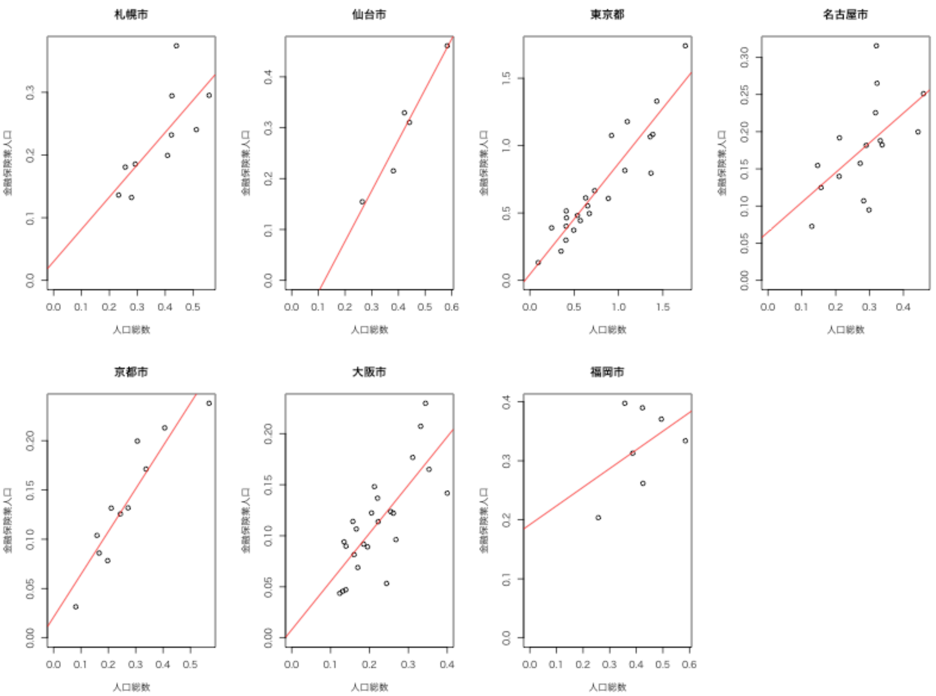

図1　最尤推定法による推定

(2)

階層ベイズ法を適用する際、Stan という MCMC サンプラーを用いて推定した。

データから σ_Y、a、b、σ_a、σ_b を推定する。Y は各年ごとのパラメータである。a, b の平均を $\hat{a}, \hat{b}$, σ_a と σ_b は a と b を決めるハイパーパラメータとする。また、$PRE[n]$ は各都市における値である。Stan で書いたモデルのモデル式は次の通り。

$$Y[n] \sim Normal(b[PRE[n]] + a[PRE[n]] \times X[n], \sigma_Y) \tag{1}$$

$$a[k] \sim Normal(\hat{a}, \sigma_a) \tag{2}$$

$$b[k] \sim Normal(\hat{b}, \sigma_b) \tag{3}$$

この書き方は、「すべての都市の平均を $\hat{a}$, であり、各都市の $a[k]$ は $\hat{a}$ を平均とした正規分布から生成された」という書き方になっている。$\hat{b}$ と $b[k]$ についても同様である。

推定の結果は次のようになった。

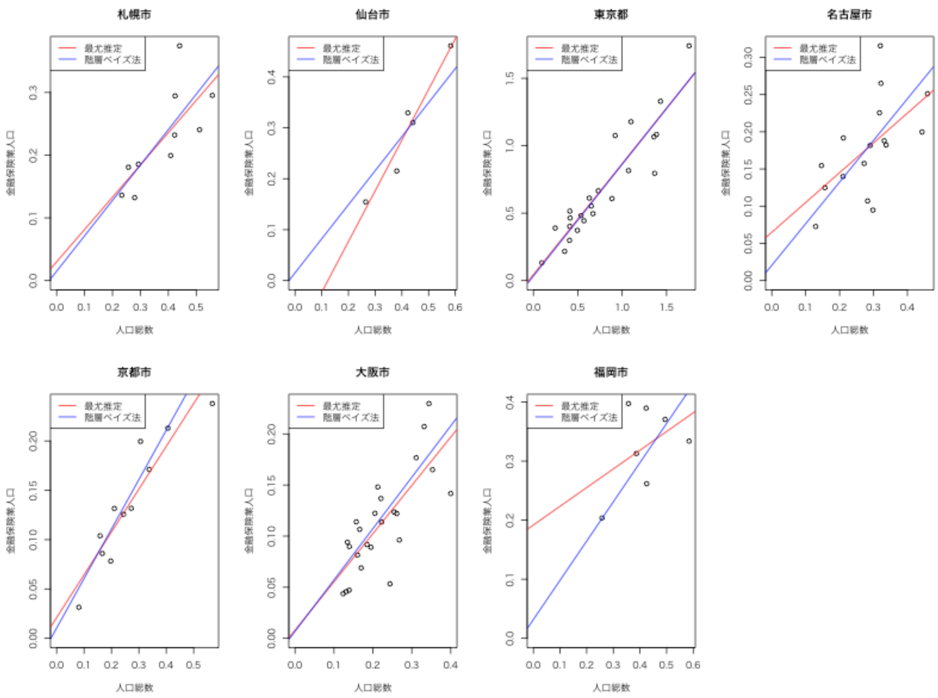

図2　階層ベイズ法と最尤推定法の比較

(3)

最も推定が大きく異なる都市は **福岡市**である。考えられる原因は以下の通り、

1. サンプル数が他の年に比べて少なく、ランダムネスによる影響が大きいため。
2. サンプルが少ない上に、点同士が反発しているような配置になっているため、よりランダムネスの影響を受けやすい（名古屋市も反発しているような配置になっているが、サンプル数が福岡市より多いため、最尤推定法と階層ベイズ法の差が福岡市に比べ小さいことより)。

また、もっとも $a(k)$ が大きい都市 k は **東京都**である。

応用問題 2