

◆ データサイエンティスト育成クラッシュコース
＜第7回＞

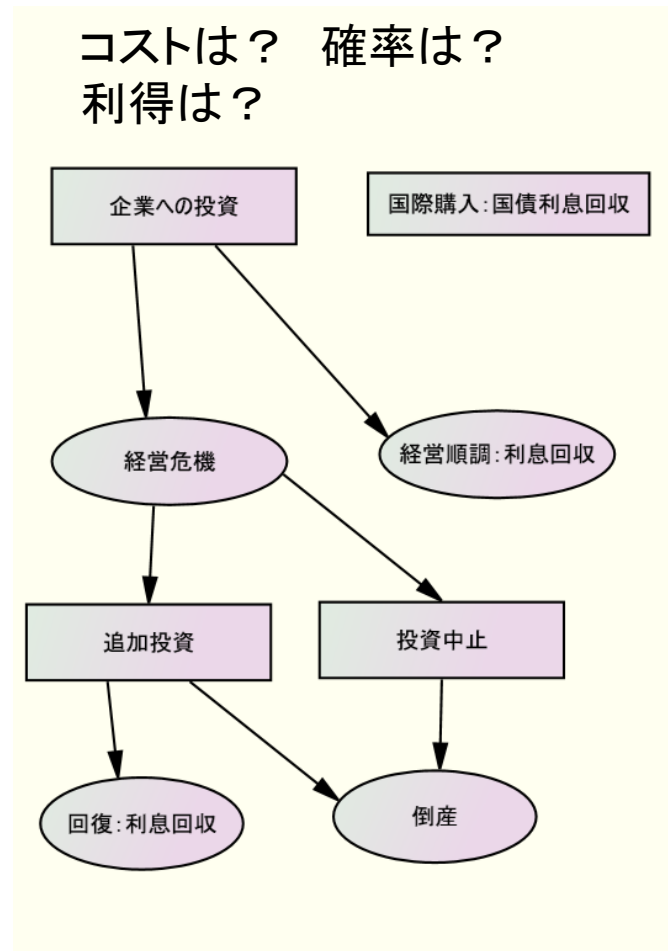
データ分析と意思決定

統計数理研究所データ科学研究系

椿 広計

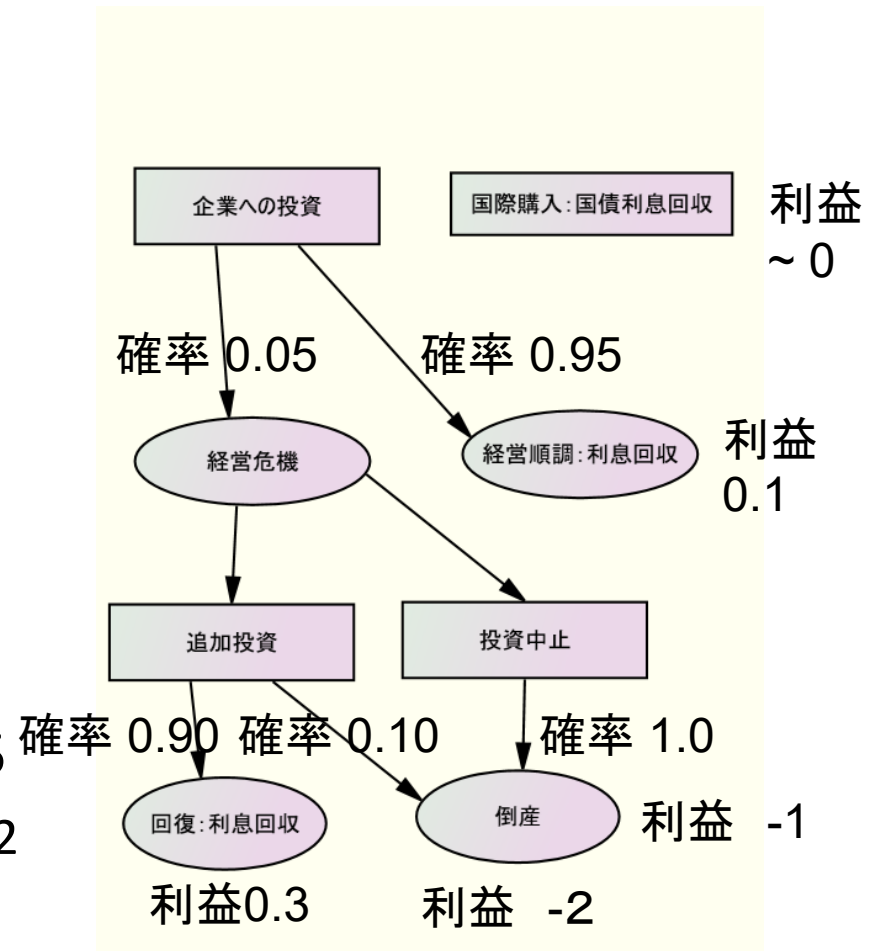
不確実性下の意思決定

- 属人的意思決定から
- 科学的意意思決定へ！
 - シナリオ・プランニング
 - 決定
 - 選択可能性のある行動の集合
 - 不確実な現象
 - 利得に影響を与える不確実な将来現象
 - » 確率でとらえるか否か
 - 利得表
 - 決定の下で、ある現象が実現した場合の利得
 - » 利得も不確実な場合が多い



期待利得最適化決定

- 後ろ向きに期待値評価
 - 追加投資の期待利得
 - $0.9 \times 0.3 + 0.1 \times -2 = 0.07$
 - 投資中止の期待利得
 - $1 \times -1 = -1$
 - 経営危機時の期待価値
 - $\text{Max}(0.07, -1) = 0.07$
 - 企業への投資の期待利得
 - $0.05 \times 0.07 + 0.95 \times 0.1 = 0.0985$
 - 経営危機確率=倒産確率=0.5
 - 企業への投資の期待利得:-0.42



社会的意思決定の難しさ

- 多様な利害関係者の存在
 - あちらを立てればこちらが立たず
- 社会全体の期待利得最適化は、特定集団に重篤な損失を与える可能性
 - 確率的制約条件が入る
 - 何らかの特性(生活水準など)が、ある値以下になってしまう確率を一定値以下に抑える
 - これにも確率予測が必要

あるべき意思決定？

- 最も保守的な行動
 - 最大損失を最小化する: マックスミニ戦略
 - 事象や損失に関する不確実性が大きいときに採用
 - この種のことが必要なのは人間がその種の情報をまだ入手していないときに起きる
 - 最大利得を最大化する: マックスマックス戦略
 - 自分の選択した行動にとって、常に都合の良い将来が実現
- 適切で計画的なデータの蓄積があれば、期待利得最適化行動やよりコストのかからないリスク回避行動が選択可能な場合がある
 - 期待利得評価モデルに基づく期待利得の最適化
 - 単純な期待利得最大化行動
 - 期待利得最大となるように決定
 - リスク回避行動
 - あまりにも大きく取り返しのつかない損失が生じる確率を一定以下に制御
 - $Prob.(大きな損失) < \alpha$ という制約を導入
- 意思決定のためのデータ解析の出番

意思決定に必要なデータ分析

- 不確実性事象が起きる確率の予測
 - 財務情報などからの倒産確率予測モデル
- 不確実な利得の予測
 - 売上高予測, コスト予測⇒利益予測
- これらのデータに基づく予測誤差が小さければ統計数理的意思決定は強力
- データに基づく統計的予測が十分可能か？
 - 予測対象の変動 = 系統変動 + 偶然変動
 - 予測値 = 系統変動, 予測誤差 = 偶然変動
 - 偶然変動が大きい
 - ビッグデータをとっても予測誤差は小さくできない
- どう対処するのか？

予測誤差が大きい場合の対処

- 多くの対象に対して繰り返し適用できる場合
 - 平均的には期待利得を稼げる
 - 大数の法則の利用
 - 1回1回の決定は、予定通りにはならない
- 将来環境の不確実な変動に対して、利得があまりばらつかない対応を行う
 - Max Min的対応より少しだけ良い決定を
 - 頑健な意思決定：転んでも大事故にはならない
 - 期待利得だけでなく、利得のバラつきも評価
 - 一定以上の損失が生じる確率の制御も有効

大数の法則の利用：

投資1期モデル→ $\Pr(\text{経営不振} | \text{予測変数}) < r / (1+r)$

無リスク投資利息 $r=0.04$ ，経営不振確率 $P < 0.038$ → 投資

- 確率予測モデルの

質は，投資機会に直結

– ロジスティック回帰
による企業不振確率予測

- 1996年上場企業2091社

- 従業員数のみ予測に利用

- » $\#(P < 0.038) = 421$

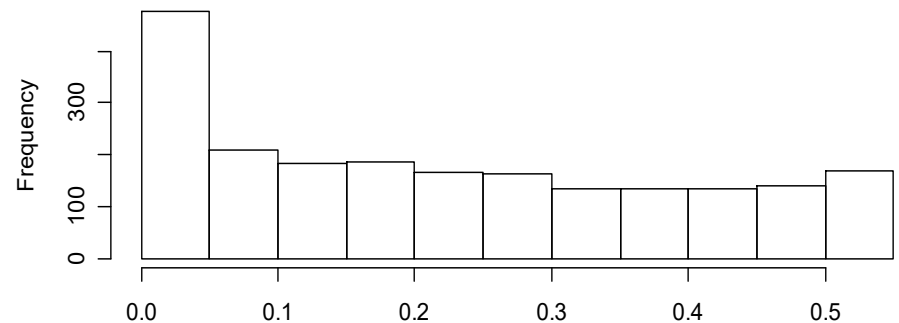
- 従業員数，資産，収入，
コスト，負債の予測の利用

- » $\#(P < 0.038) = 1394$

- » 予測式導出によって
投資機会3倍

- 椿(2006)ビジネスへの統計モデルアプローチ，
シリーズビジネスの数理4，朝倉書店。

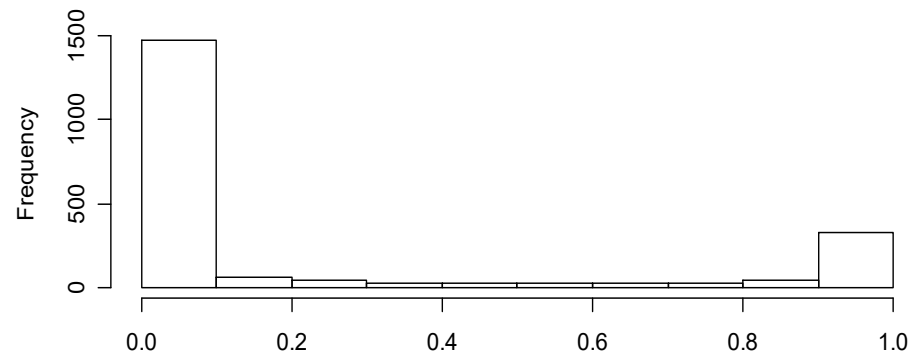
Histogram of fitted(safe.glm1)



fitted(safe.glm1)

企業の安全性がC,CCとなる予測確率

Histogram of fitted(safe.glm2)



文献

- 不確実性下のマネジメント, 情報の価値入門
 - Bernard W. Taylor (2012), *Introduction to Management Science*, 11th ed., Pearson Education.
 - 日本にこの種の初等テキストが無いのが残念
- 回帰分析を少し超えた予測モデルなどのデータ分析(Rなどによる)入門: 習うより慣れろ
 - ヴェナブルズ, リプリー, 伊藤訳(2012), S-Plusによる統計解析第2版, 丸善. 6章から9章
 - Oxford大学統計学科のデータ解析教科書
 - 2000年ころまでの方法をRなどで習得し総覧するには便利