リスク評価におけるベイズ手法活用について（2）ベイズ統計による信頼性パラメータ評価

日本原子力学会　2019秋の大会

2019年9月11日　リスク部会セッション

吉田智朗

電力中央研究所　原子力リスク研究センター

tomo@criepi.denken.or.jp

リスク評価と不確かさ

- 確率論的リスク評価では、不確かさを確率分布で表現する。
- 不確かさとその確率表現には2種類がある。

•偶然の不確かさ (aleatory uncertainty)

確率＝無限反復試行の**極限頻度** (frequency)
事象(＝データ)が確率過程に従って**ランダムに**発生するため、
次の試行でどういう**データ**が得られるか不確定。

•認識の不確かさ (epistemic uncertainty)

確率＝命題の真偽に対する**確信の度合い** (degrees of belief)
命題が真なのかどうか不確定。

確率論的リスク評価では、特に認識の不確かさが重要

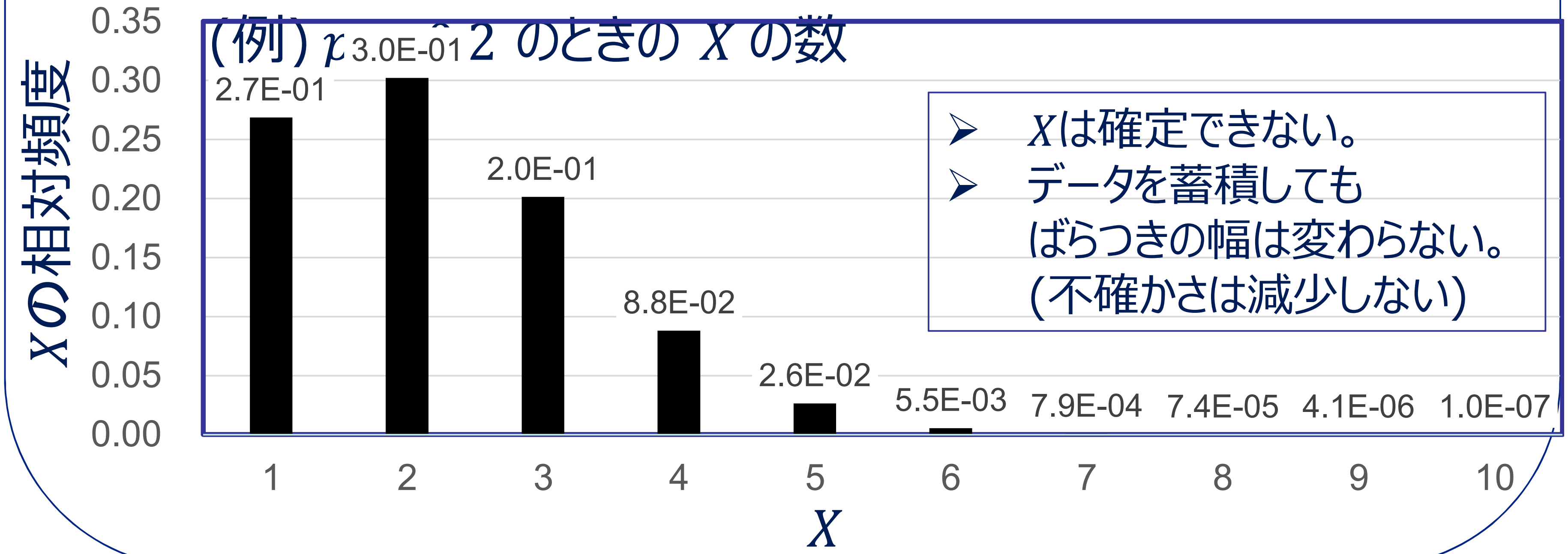
偶然の不確かさ aleatory uncertainty

- 10個の同じ(いびつな)サイコロを投げたとき、**1の目の出る数** X がいくつになるか**確定できない。**

$X = 0,1,2,\cdots,10$
$0 < p < 1$

  - 上のサイコロを無限回投げたときの1の目が出る数の割合をpとすると、
    X は二項過程 $P(X) = {}_{10}C_X p^X (1-p)^{10-X}$ に従う。

(例) $p = 0.2$ のときの X の数



- ➢ Xは確定できない。
- ➢ データを蓄積しても
ばらつきの幅は変わらない。
(不確かさは減少しない)

認識の不確かさ epistemic uncertainty

前ページの確率過程の例のように、

$$P(X) = {}_{10}C_X p^X (1-p)^{10-X}$$

実際に得られるのは1の目の出る数Xの観測データであって、

pの値がいくつなのかは確定できない ⇒ 認識の不確かさ

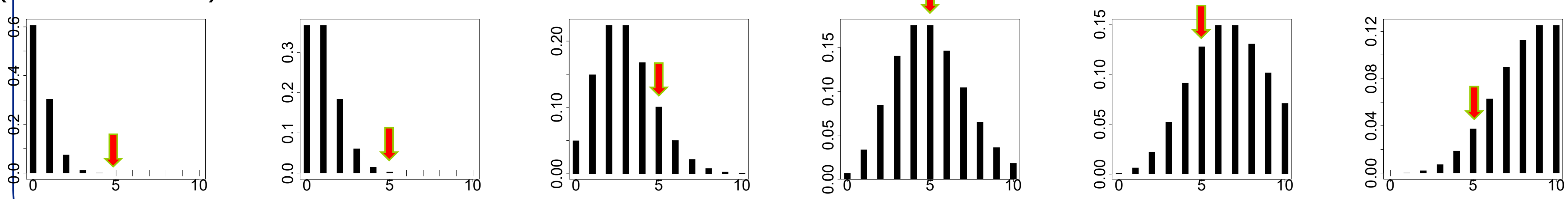
母数の不確かさ

【例】故障確率 p に対して下記観測データがえられたとする。

観測データ：起動デマンド 10000回、起動失敗 5回

pの不確かさ：起動失敗5回を生じさせるpは無数にある。

$p=\dfrac{1}{20000}$	$p=\dfrac{1}{10000}$	$p=\dfrac{3}{10000}$	$p=\dfrac{5}{10000}$	$p=\dfrac{7}{10000}$	$p=\dfrac{1}{1000}$
1.6×10^{-4} (5回発生する確率)	3.1×10^{-3}	0.1	0.18	0.13	3.8×10^{-2}



起動失敗確率 p のときの、起動デマンド 10000回あたり起動失敗数の分布

PRAにおける偶然と認識の不確かさの区別(分離?)

- 偶然の不確かさ
  各種事象(機器故障・起因事象・ハザード、待機除外時間、etc.)がランダムに発生して予測できないこと
- 偶然の不確かさの考慮
  上記事象をランダム事象と認識して確率過程モデルで表現すること
- 認識の不確かさ
  上記確率過程モデルの分布形の選択、確率分布母数が確定できないこと。
- 認識の不確かさの考慮
  上記分布形、母数の推定を行うこと。

統計的推測

- 頻度論統計手法は、確率を極限頻度で定義しているため偶然の不確かさしか扱えず、認識の不確かさを扱うことができない。(pの不確かさを表す確率分布を求めることができない)

- ベイズ統計手法は、確率解釈を問わないベイズの定理を用いるため、**偶然・認識いずれの不確かさも扱うことができる。**
 (pの不確かさを表す確率分布を求めることができる)

ベイズの定理

- 確率変数 ϕ, D　（確率解釈は問わない）
- 条件付き確率の式を変形しただけ。

$$P(\phi, D)$$
$$= P(D)P(\phi|D) = P(\phi)P(D|\phi)$$

$$P(\phi|D) = \frac{\boldsymbol{P(\phi)P(D|\phi)}}{P(D)} = \frac{\boldsymbol{P(\phi)P(D|\phi)}}{\underbrace{\int_{\phi'} \boldsymbol{P(\phi')P(D|\phi')}\, d\phi'}_{\text{規格化定数}}}$$

$$\propto \boldsymbol{P(\phi)P(D|\phi)}$$

ベイズの定理による母数推定

パラメータ、例：前ページのpやλ

例：前ページの二項過程やポアソン過程

■ **母数**ϕで特徴づけられる**確率過程**において、データDを得て母数ϕを推定する場合

$$P(\phi|D) \propto P(\phi)P(D|\phi)$$

データDを得たあとの母数ϕの**事後分布** - epistemic

データDを得る前の母数ϕの**事前分布** - epistemic

「母数ϕで特徴づけられる確率過程によりデータDが出現する確率」 - aleatory をϕの関数とみたもの－**尤度関数** という

データD以外の情報 (過去・他所の評価，専門家判断など) を母数ϕの推定に使う。

↓

「データが少なくても推定が可能」の意

データDからわかる母数ϕの範囲を示す。

（頻度論統計の推定法では事前分布は使わず、**尤度関数**$P(D|\phi)$**を最大とする**ϕの値をϕの点推定値とする。【最尤法】）

デマンド故障確率の例

- 待機機器に起動デマンドをかけたときに起動失敗する確率、デマンド故障確率を p とする。
- データ：起動デマンド数N、それに対する起動失敗回数k
- 失敗回数kの従う確率過程モデルは二項過程
- データkを得たあとのpの事後分布

$$P(p|k,N) \propto P(p)\cdot p^k(1-p)^{N-k}$$

【例(左図)】

事前分布 (ベータ分布・・・二項分布と共役)

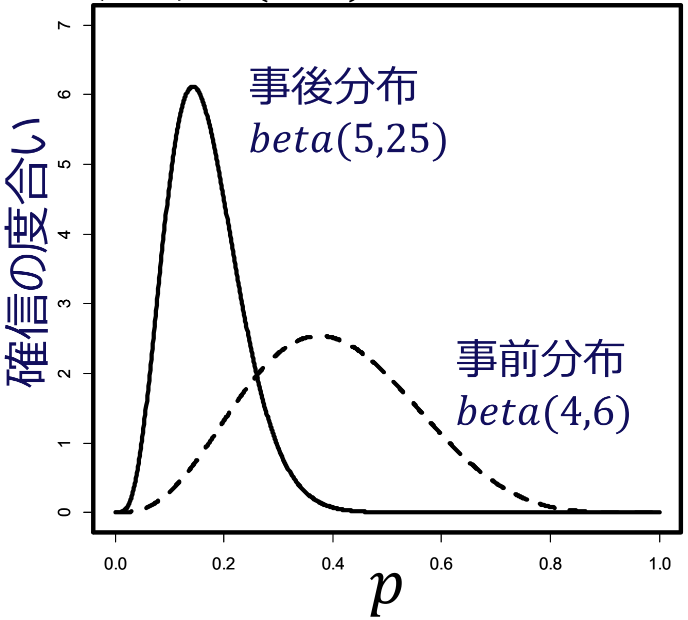
$$P(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}$$

事後分布

$$P(p|k,N) = \frac{\Gamma(\alpha+\beta+N)}{\Gamma(\alpha+k)\Gamma(\beta+N-k)}p^{\alpha+k-1}(1-p)^{\beta+N-k-1}$$

$(\alpha,\beta) = (4,6)$　事前分布形状

$(k,N) = (1,20)$ 失敗1/20デマンド



その他各種信頼性パラメータ

- 時間故障率
  - データ：観測時間内の故障件数
  - 確率過程：ある時間内にある割合で故障がポアソン過程でランダムに発生する
- 待機安全系継続運転時間
  - データ：継続運転失敗時間、成功打ち切り時間
  - 確率過程：継続運転失敗の時間がワイブル過程でランダムに発生する
- 共通原因故障
  - データ：(単独故障件数,2機同時故障件数,3機同時故障件数,…)
  - 確率過程：単独、2機同時、3機同時の故障件数が、ディリクレ分布に従い、それぞれある比率で発生する [アルファファクタの場合]

階層ベイズモデルの概要

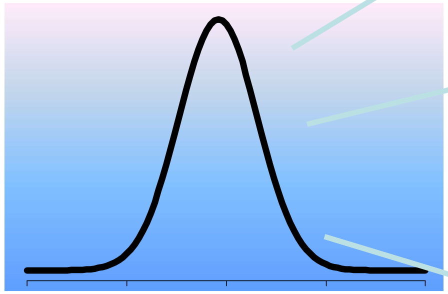
国内個別故障確率$p_1, p_2, \dots, p_M$が全体であるばらつき分布をなすという状況

国内故障確率分布（母集団ばらつき分布）

$p = p(\theta)$

超母数θ

（ハイパーパラメータ）



$\boldsymbol{p}$

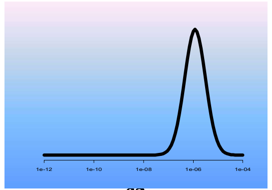
二項過程の母数が$p_{1\sim M}$

pのばらつき分布の母数がθ

個別プラント故障確率（不確実さ分布）

p_1

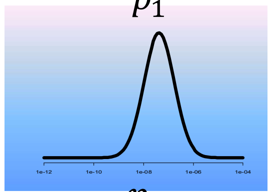
プラント1



p_1

p_1

プラント2

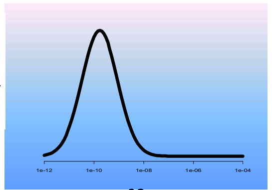


p_2

・
・
・

p_M

プラントM



p_M

個別プラントデータ

二項過程で発生 → (k_1, N_1)

二項過程で発生 → (k_2, N_2)

・
・
・

二項過程で発生 → (k_M, N_M)

※計算にはマルコフ連鎖モンテカルロ法を用いることがきわめて多い。

国内ベイズ信頼性パラメータ推定の課題1/3

■ 極少故障件数の扱いと事前分布選択

- 極小件数ゆえ、従来計算していたの国内パラメータ(JANTI,JANSI)は概して海外水準に比べ1~2桁低かった。海外から疑問の声あるも、妥当性が説明できる確実なデータはなかった(特に機器母集団、exposure data)。
※ベイズ手法の問題ではない

【対応】

プラント個別で必要な母集団、exposureデータを収集しつつあり、データの妥当性を証明できるようにする。

国内ベイズ信頼性パラメータ推定の課題2/3

■ 極少故障件数の扱いと事前分布選択

- 極小件数ゆえ、推定には事前知識が必要(ベイズ手法の特長)だが、適切と思われる事前知識がない。海外データ(特に米国)を事前知識とすると、一般的に数値が高くなる傾向にあり、国内故障件数の実績水準と整合しない場合がある。

【対応】

経験ベイズ(実績データを事前分布に利用する)を行う、あるいは、一部米国データを事前知識利用する。
(高い故障率の数値は今後の実績データ蓄積で更新すれば適正な水準になるはず)

国内ベイズ信頼性パラメータ推定の課題3/3

■ プラント個別評価手法の適切性

- 従来は、プラント個別評価をするため階層ベイズ手法を採用したが、極少件数のためモンテカルロ計算の収束が悪い。特に、プラント間ばらつきのみられないデータに階層ベイズを用いるのは不適切。

【対応】

データから検定等により一次的なプラント別ばらつきを判断し、ばらつきのみられないものには階層ベイズは使用しない。米国の一般パラメータ推定手法も参考にして、故障件数データの少ないものに大してはなるべくシンプルな手法(前ページ)を適用する。