

組織の属性に基づくインシデント規模と頻度のモデル提案

2021/03/01(月)

ICSS 2021

池上和輝 菊池浩明

明治大学大学院

背景

■ 内部犯行、不正アクセスによる情報漏洩(インシデント)増加

- 2020年12月, PayPayの不正アクセス被害
- 2018年 443件, 561万件の個人情報流出[1]

■ 企業のリスク認識の課題

- 企業が被害に遭う確率を過少に見積もっている
- ISMS認証などのコストに対して, それに見合う必要性を認識していない
- セキュリティ保険加入率も諸外国より低い[2]
 - ・ 情報漏洩のリスクを感じない
 - ・ 費用対効果がわからない

■ セキュリティリスクの簡易的な定量化が必要



[1]日本ネットワークセキュリティ協会, 2018年 情報セキュリティインシデントに関する調査報告書~個人情報漏えい編~(速報版).

[2]佐久間樹里, 猪俣敦夫, サイバー保険の調査・分析による加入率向上への提案, 研究報告インターネットと運用技術(IOT)(IPSJ), pp. 1-8, 2019

先行研究と問題点

■ 先行研究

	Edwardsら[3]	山田ら[4]	本研究
研究目的	米インシデントのトレンド調査	マネジメント方策によるインシデント削減効果	
対象	アメリカの全企業	日本の（単一）企業	
目的変数	被害規模, 頻度	生起確率	
説明変数	時間（日）	マネジメント方策, 業種, 企業規模	
結果	2005-2015年でトレンド変化なし	CIO設置により, インシデントを0.3倍に抑制	

■ 問題点

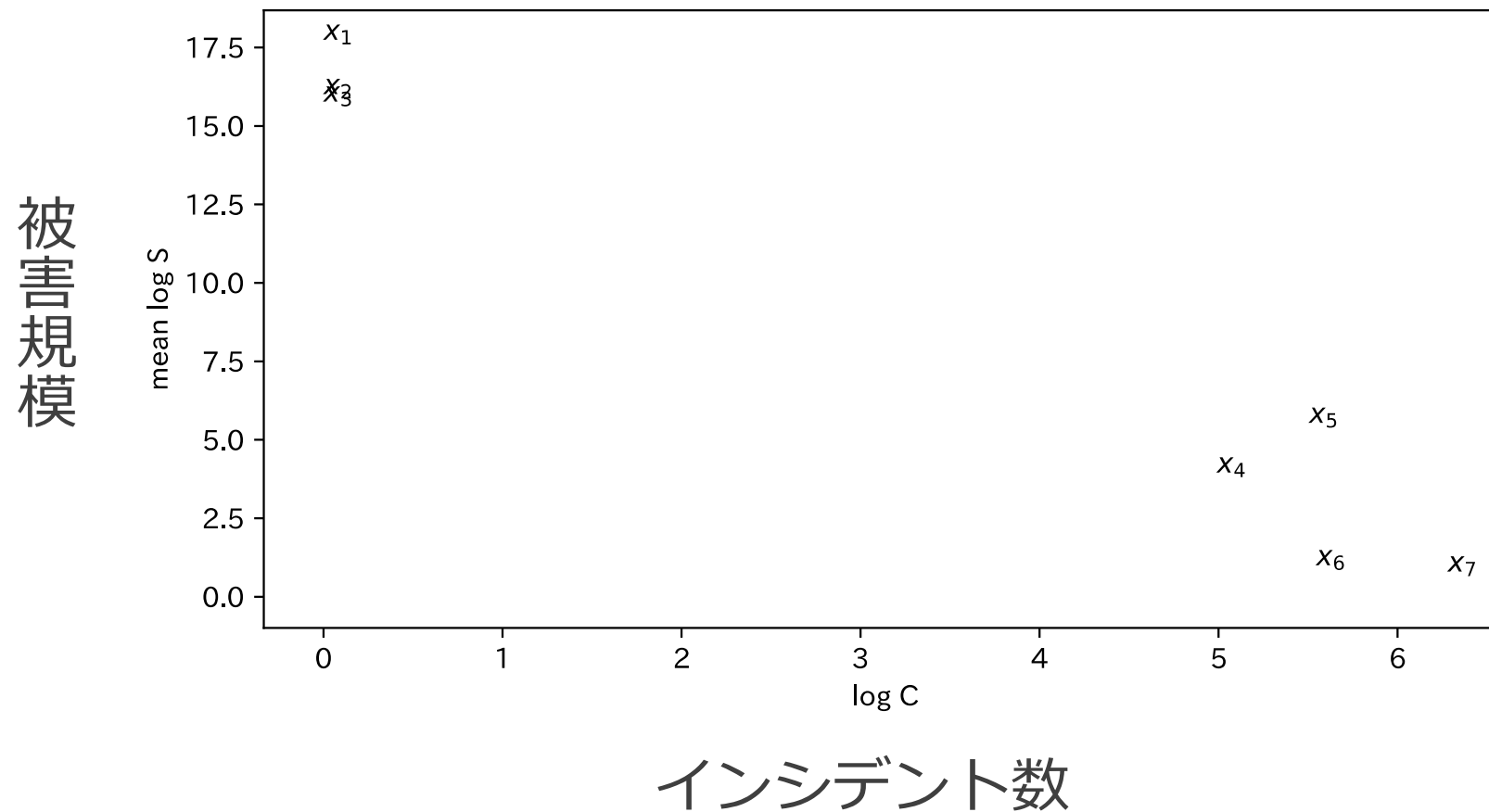
- 被害規模と頻度を独立にモデル化していること
➡ 被害規模ごとの頻度が不明確
- 業種など他の要素を考慮していないこと
➡ 振る舞いの異なるインシデントを混ぜると正しくリスクを測れない

[3] B.~Edwards, S.~Hofmeyr, and S.~Forrest, Hype and heavy tails: A closer look at data breaches, Journal of Cybersecurity, 2(1): 3-14, 2016.

[4] 山田道洋, 池上和輝, 菊池浩明, 乾考治, 経営マネジメント状況による情報漏洩インシデント削減効果の評価(2), CSS2018

被害規模とインシデント数

- 2005 – 2018年, 9,007組織のインシデント数と平均被害人数
- 被害規模とインシデント数に**負の相関**



研究目的

- 目的：将来起こるインシデントの被害規模と頻度を算出すること
 - 被害規模と頻度を影響し合う因子として扱う
 - 業種，漏洩原因を分けてモデルを作成

	Edwardsら[3]	山田ら[4]	本研究
研究目的	米インシデントのトレンド調査	マネジメント方策効果	被害規模と頻度の算出
対象	アメリカの全企業	日本の（単一）企業	日本の（単一）企業
目的変数	被害規模， 頻度	生起確率	頻度
説明変数	時間（日）	マネジメント方策， 業種， 企業規模	被害規模， 業種， 漏洩原因

[3] B.~Edwards, S.~Hofmeyr, and S.~Forrest, Hype and heavy tails: A closer look at data breaches, Journal of Cybersecurity, 2(1):3--14, 2016.

[4] 山田道洋, 池上和輝, 菊池浩明, 乾考治, 経営マネジメント状況による情報漏洩インシデント削減効果の評価(2), CSS2018

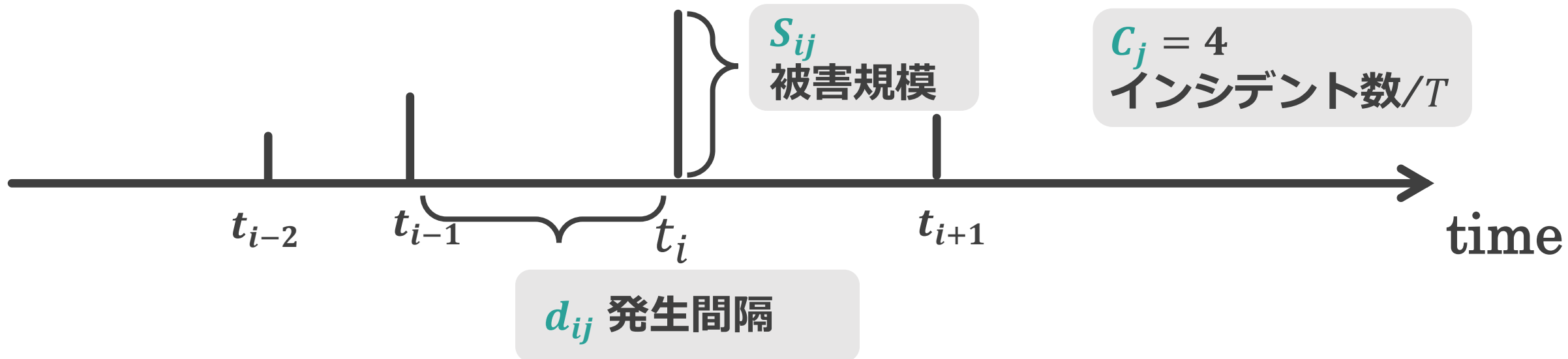
データ

■ JNSAデータセット (2005-2018)

- 新聞やインターネット, 企業のリリース等の公開情報から収集
- 9,007組織の15,604インシデント使用

■ 使用するインシデント情報

- T 年間で組織 j が i 番目に起こしたインシデントから取得する情報

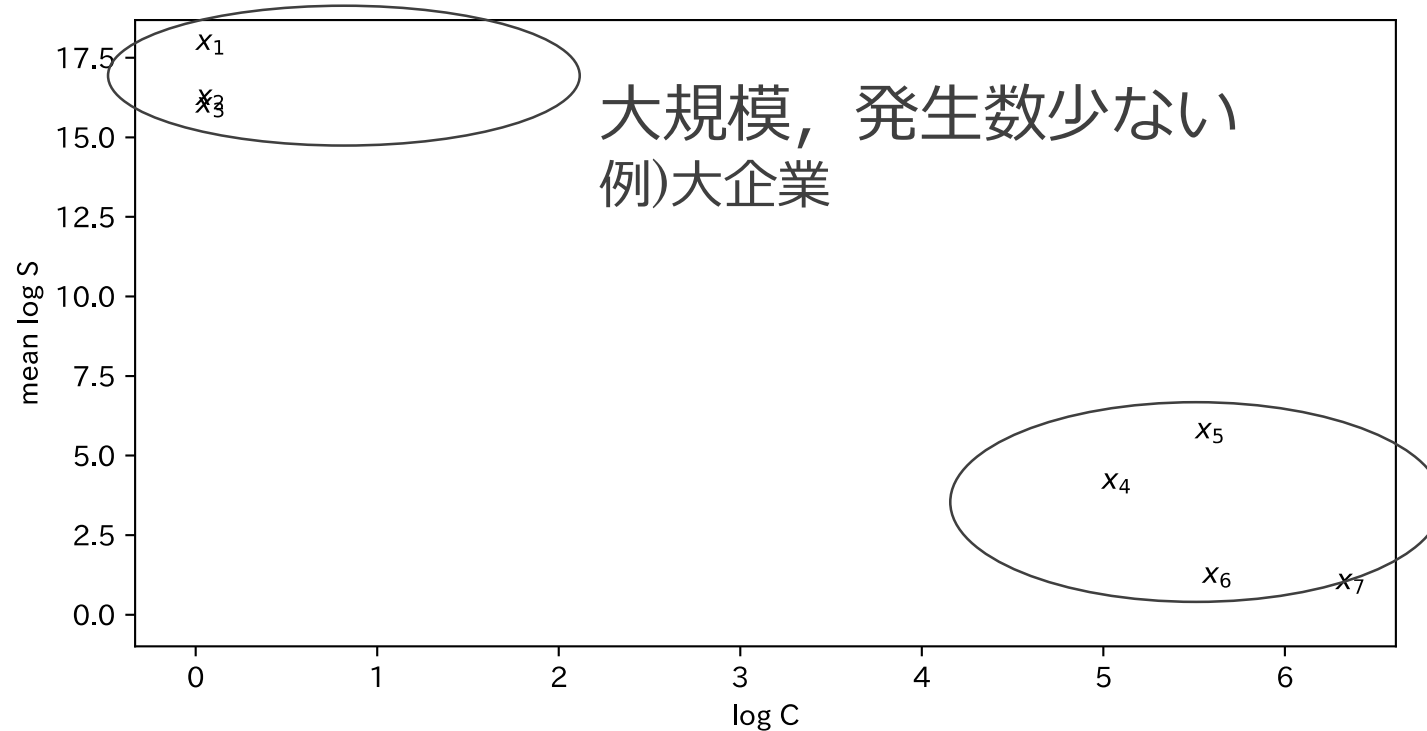


分析の流れ

1. 被害規模 S を複数の視点から分析
 - ① 被害規模 S とインシデント数 C
 - ② 企業ごとの被害規模 S の分布
 - ③ 漏洩原因ごとの被害規模 S の分布
2. モデルの提案
3. モデルの評価

分析①：被害規模 S とインシデント数 C

- 9,007企業のインシデント数 C と平均被害人数 S
- 被害規模とインシデント数に負の相関



- 仮説) 被害規模は企業によって固有に決まる

分析②：企業と被害規模の独立性

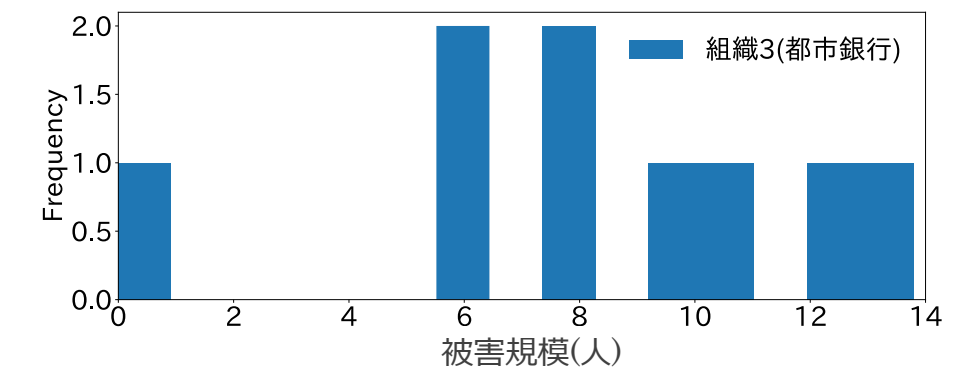
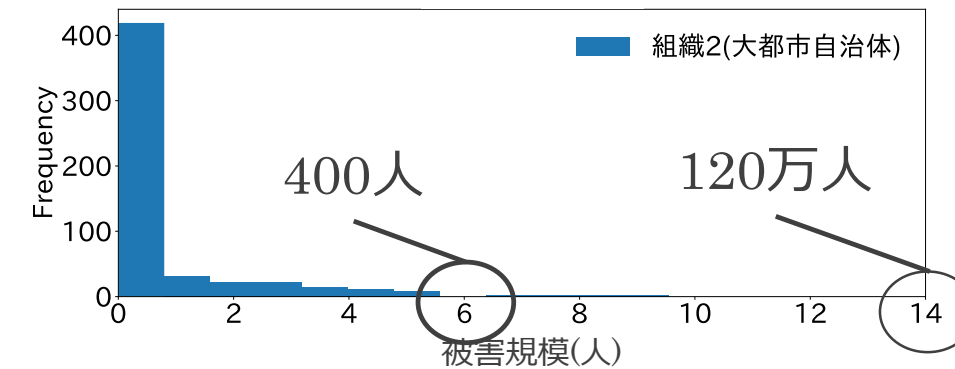
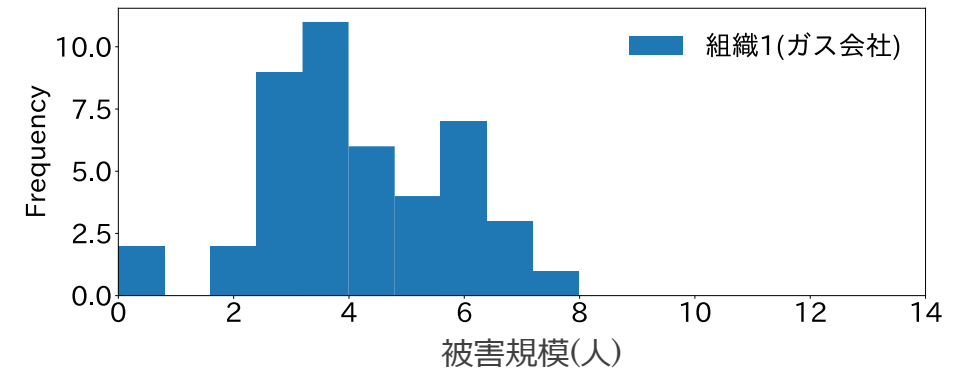
■ 仮説

- A) 被害規模は企業によって固有に決まる
- B) 被害規模と企業によって定まらない
(小規模も大規模も起こる)

■ 主要な企業の被害規模の分布

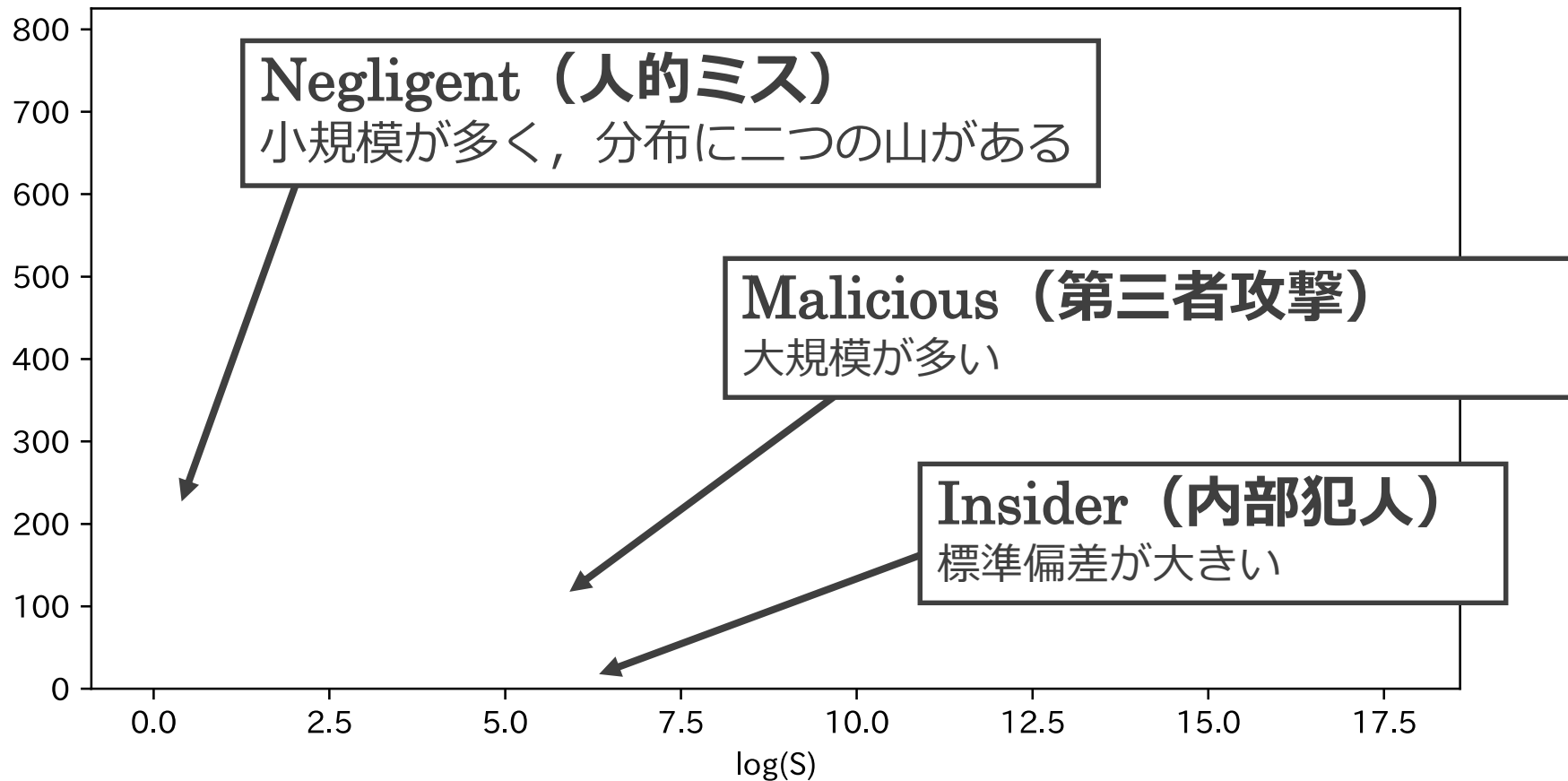
	分布	仮説Aの成立
組織1	正規分布	×
組織2	小規模に偏り	○
組織3	大規模に偏り	×

- 一部の組織で、仮説Aが成り立つが多くの場合で**組織と被害規模は独立**



分析③：漏洩原因ごとのSの分布

- JNSAの11漏洩原因を3種類に分類



- 3種類の漏洩原因によって被害人数が異なる

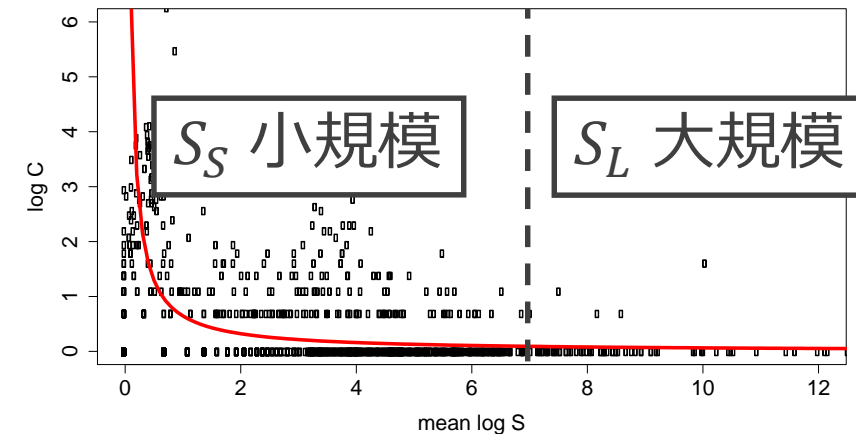
提案モデル

■モデル作成

- インシデント数 C と被害人数 S に負の相関あり (分析①)

$$\cdot \log C_j = \frac{1}{\alpha \log S_j}$$

- $\alpha_{k\ell}$ は, 3原因 k と16業種 ℓ から推定
 - ・ 漏洩原因ごとに分布が異なる (分析③)
 - ・ 業種がインシデント生起確率に影響[2]



■モデル利用, 評価

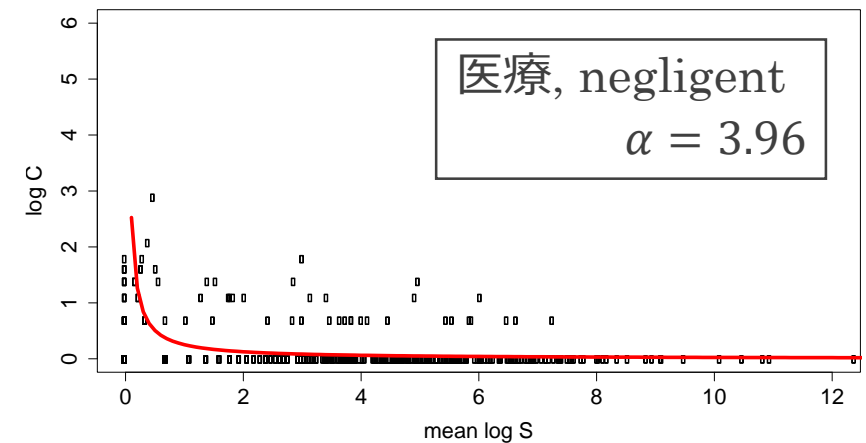
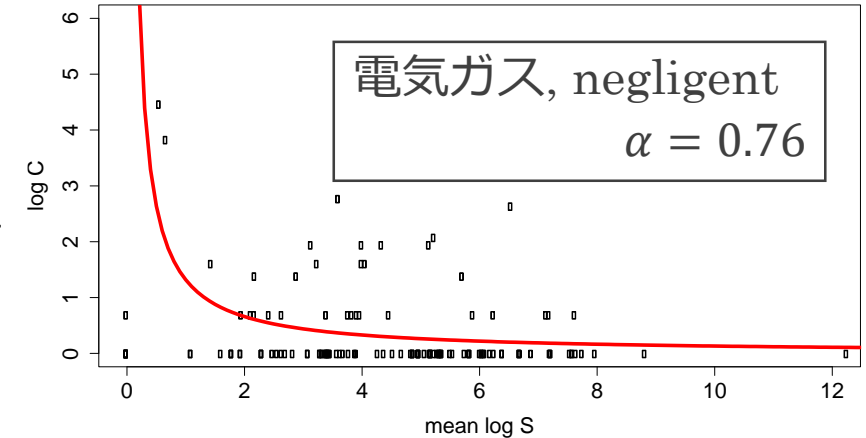
- 各モデルで小規模 S_S と大規模 S_L のインシデント数 C を予測
 - ・ 被害規模は組織によって定まらない (分析②)

被害規模	分類方法
S_S	$S_j < 1,000$
S_L	$S_j \geq 1,000$

推定パラメータ (一部)

- 46種類(3原因, 16業種)のパラメータ α の一部
- α は安全度示す

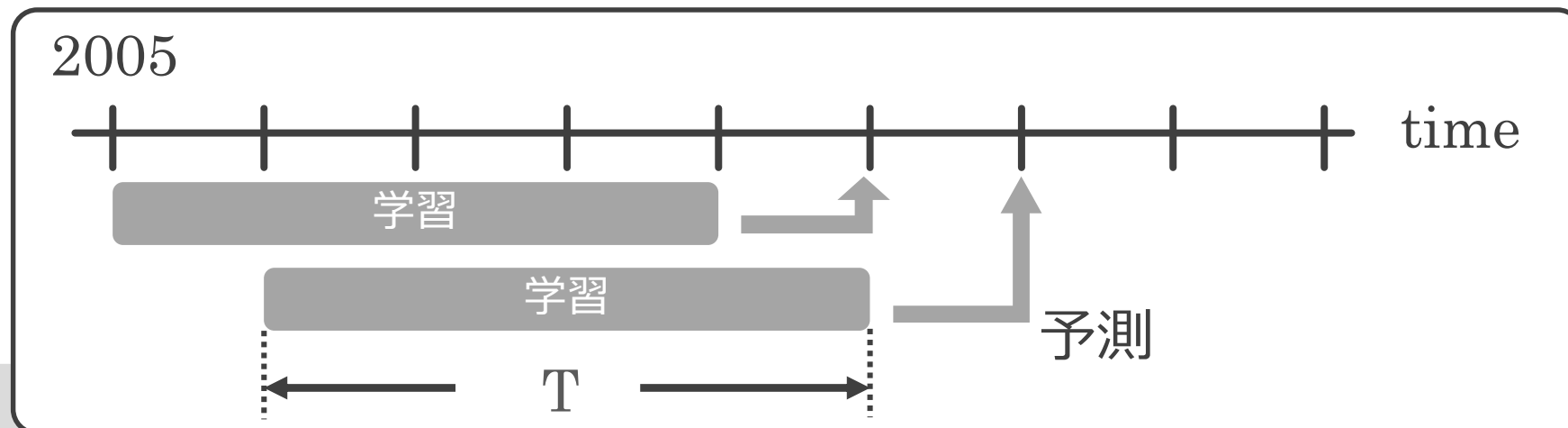
業種 k	漏洩原因 ℓ		
	negligent	malicious	insider
電気ガス業	0.76	1.01	5.81
金融業, 保険業	1.24	3.00	2.79
建設業	1.43	2.08	1.67
⋮	⋮	⋮	⋮
医療, 福祉	3.96	6.01	2.66
宿泊業, 飲食業	4.89	2.74	1.40E+09
サービス業	5.80	5.38	9.64



モデル評価手順

1. 2005年から $T = 5$ 年間を学習データとして α 推定
2010($2005 + T + 1$)年をテストデータ
2. 推定：大規模($S_L = 1,000$), 小規模($S_S = 1.5$) $\hat{C}_{k\ell}(S_L)$, $\hat{C}_{k\ell}(S_S)$
3. 観測：テストデータ内で組織 j 毎に規模 S の C 集計 $C_{j\ell}(S_L)$, $C_{j\ell}(S_S)$
4. 組織 j , 規模 S , 原因 ℓ , j の属する業種 k 毎に誤差算出

$$E_{j\ell}(S) = \exp \left| \frac{\log \hat{C}_{k\ell}(S)}{T} - \log C_{j\ell}(S) \right|$$



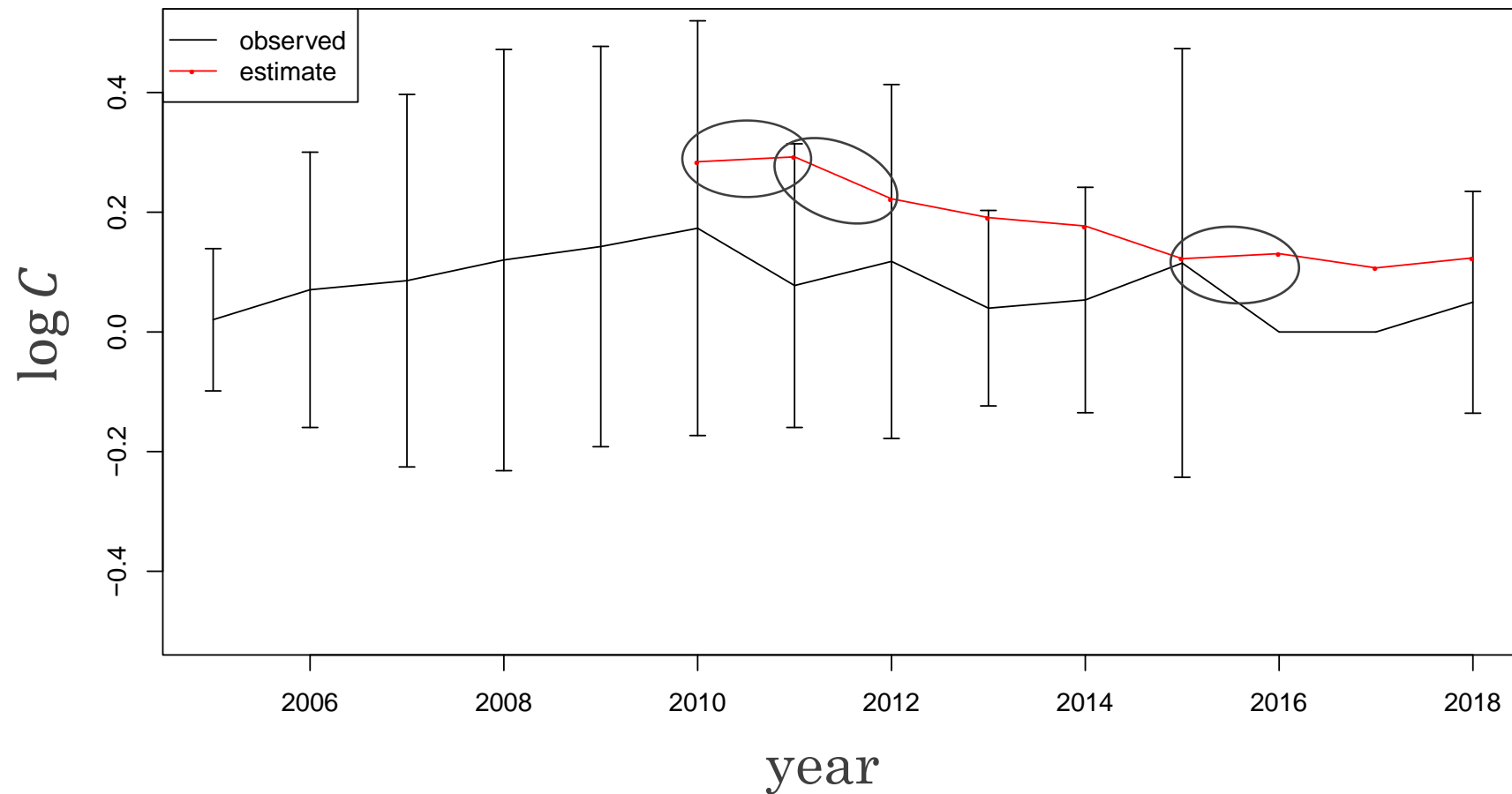
評価結果(negligent, S_S)

業種 k	\bar{c}	$\overline{\log S}$	$E(\text{negligent}, S_S)$	
			平均	最大
電気・ガス・熱供給・水道業	2.94	3.29	1.8	12.7
公務	2.95	1.92	1.6	127.7
金融業, 保険業	1.80	5.49	1.4	114.4
情報通信業	1.47	4.84	1.2	4.4
医療, 福祉	1.32	3.38	1.1	7.5
学術研究, 専門サービス業	1.06	5.28	1.1	1.9
⋮	⋮	⋮	⋮	⋮
全データ使用	1.73	4.66	1.3	130.3

- 平均誤差最大の業種は、電気ガス業で1.8件
- 1企業の最大誤差は、公務127件
- 全データを使用した場合と比較して平均誤差が0.76倍に改善

予測値と実測値の分布

- 情報通信業, negligent, S_S
- 平均誤差1.23件



まとめ

- インシデント数と被害人数との相関から、被害人数についてインシデント発生数を定量化するモデルを提案

- $\log C_j = \frac{1}{\alpha \log S_j}$

- 企業によって固有の被害人数が定まらない
- 3漏洩, 16業種をもとに48モデルを作成, 被害規模ごとに評価
 - Negligent, 小規模 S_S
 - ・ 平均誤差最大の業種は, 電気ガス業で1.8件
 - ・ 1企業の最大誤差は, 公務で127件
 - ・ 全データを使用した場合と比較して平均誤差が0.76倍に改善
 - Malicious, 大規模 S_L , 平均1.00件の誤差
 - Insider, 小規模 S_S , 平均1.05件の誤差

質疑応答用

評価結果詳細

業種名	mean(C)	logS	negligent, S_S	malicious, S_L	Insider, S_S
電気・ガス・熱供給・水道業	2.94	3.29	1.76	1.02	1.01
公務(他に分類されるものを除く)	2.95	1.92	1.56	1.02	1.10
金融業, 保険業	1.80	5.49	1.41	1.02	1.11
複合サービス事業	1.30	5.11	1.31	1.00	1.00
不動産業, 物品賃貸業	1.68	3.16	1.28	1.01	1.00
サービス業(他に分類されないもの)	1.12	5.43	1.28	1.00	-
運輸業, 郵便業	1.35	5.65	1.26	1.02	1.00
建設業	1.66	3.93	1.25	1.01	1.00
情報通信業	1.47	4.84	1.24	1.02	1.38
卸売業, 小売業	1.21	5.54	1.19	1.00	1.00
教育, 学習支援業	1.28	4.35	1.17	1.01	1.05
生活関連サービス業, 娯楽業	1.10	5.60	1.14	1.00	1.00
医療, 福祉	1.32	3.38	1.13	1.00	1.12
製造業	1.17	5.95	1.13	1.00	1.00
宿泊業, 飲食サービス業	1.11	5.67	1.07	1.01	-
学術研究, 専門・技術サービス業	1.06	5.28	1.05	1.00	-