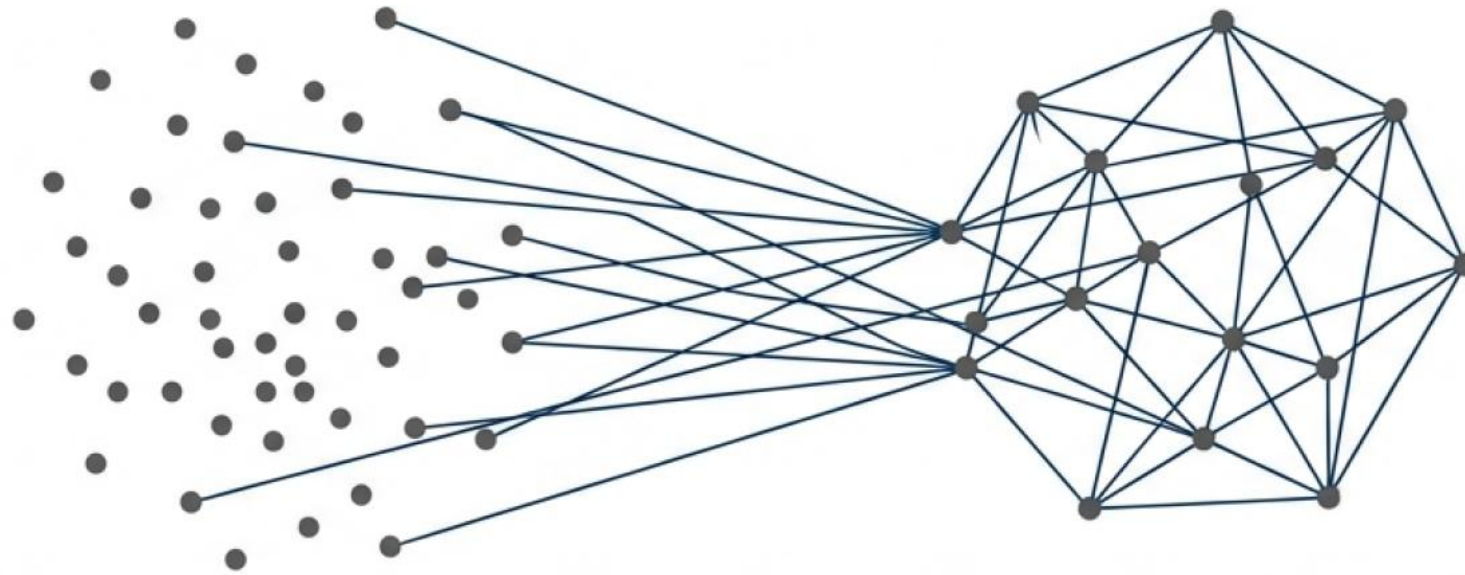


アソシエーション分析に基づく 再検査基準の決定手法の提案



第41年度（2025年度）ソフトウェアテストコース Team GATE
リーダー： 片桐 汐駿 （アズビル株式会社）
研究員： 若林 俊弥（株式会社日立システムズ）
越智 恭平 （アズビル株式会社）
荻原 美樹 （アズビル株式会社）

Question!

品質基準って
あいまいじゃないですか？

Question!

どうしてその基準に
なっているのか説明できますか？

Question!

業界の平均

経験則による
区切りの良いしきい値

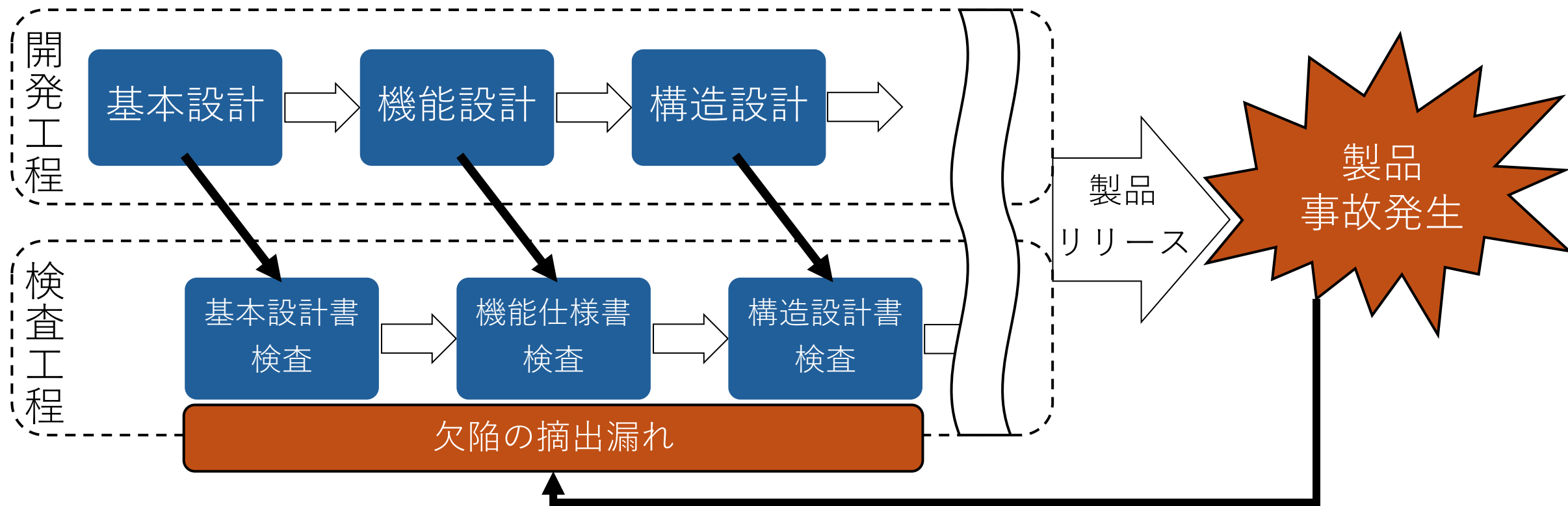
どうしてその基準に
なっているのか説明できますか？

根拠のないしきい値

ネットに
書いてある基準

欠陥が検査をすり抜ける

各開発工程でドキュメント検査※合格後に重大な欠陥が見逃されるケースが存在



※ドキュメント検査：設計仕様書の内容を第三者が客観的に評価する

検査の指標値に問題があるのでは？

設計品質を確保し、工程完了を判断する必要があるが、基準を満たしていても重大な欠陥が見逃され、後工程や製品リリース後に事故が発生するケースが存在



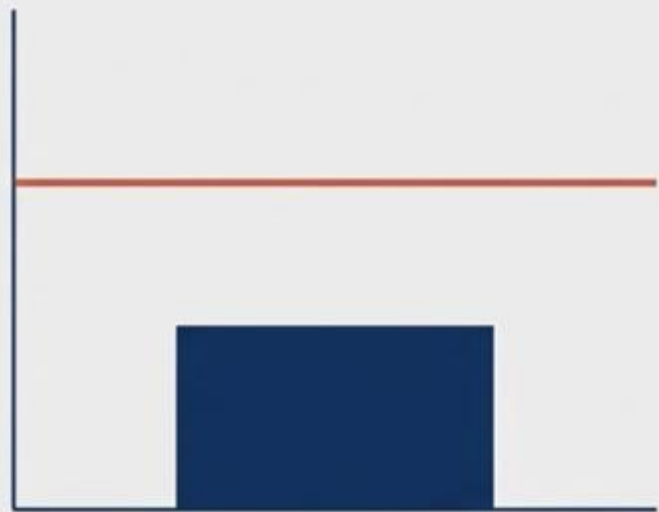
不良密度※を指標値として、
5.0未満であれば品質良と判断

しきい値は経験則で定められている
指標値の妥当性を説明できない
全プロジェクトで同じ

※不良密度：設計書100ページ当たりの欠陥数

それ、GATEが一つの解となるかもしれません

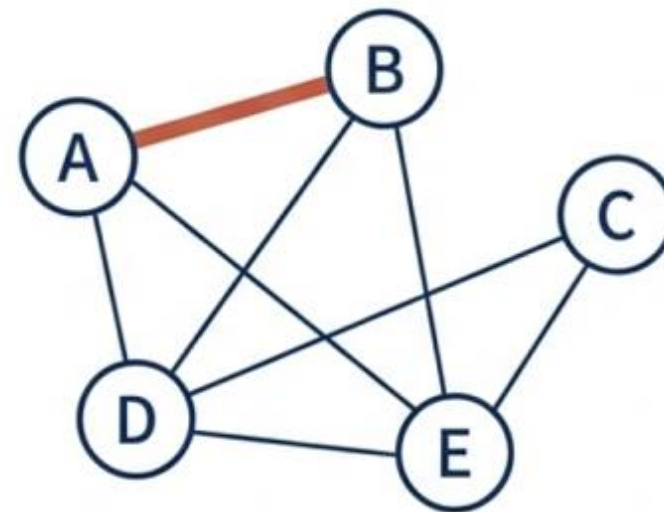
従来の手法



不良密度を評価する

GATE

Guidelines by Association Threshold Evaluation



アソシエーション分析
Association Analysis

検査後の潜在的な欠陥を推定し、再検査すべき箇所と併せて再検査を指示する

こんな経験ありませんか？



欠陥A

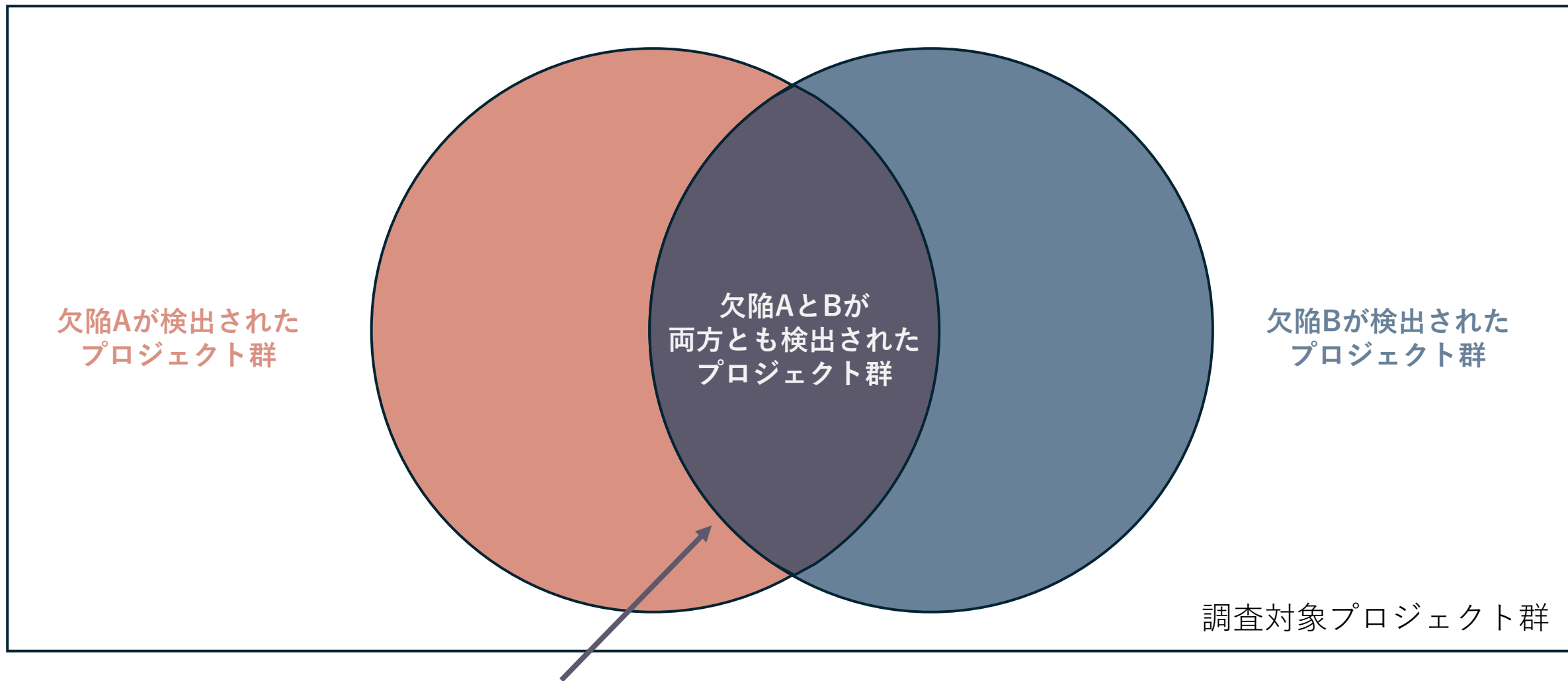


欠陥B

「欠陥Aがあるソフトウェアには、欠陥Bも潜んでいる」

欠陥は単独ではなく、群れで発生する

欠陥Aの検出をトリガーとして、欠陥Bを効率的に探索するためには



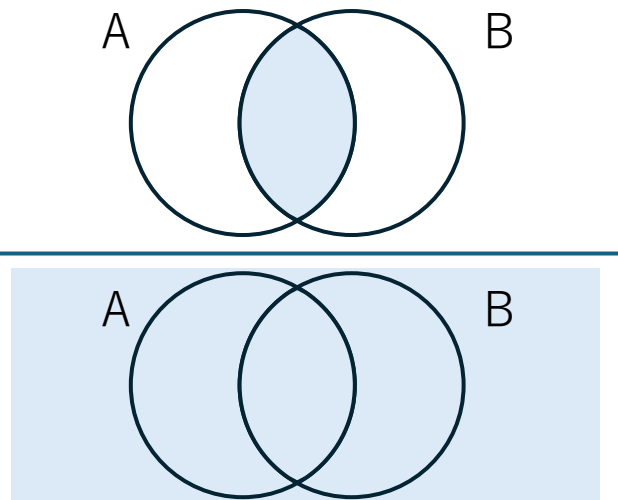
欠陥AとBが両方検出されたプロジェクト群に注目する

意思決定のカギとなる指標

支持度

AとBが同じプロジェクトで起きる確率

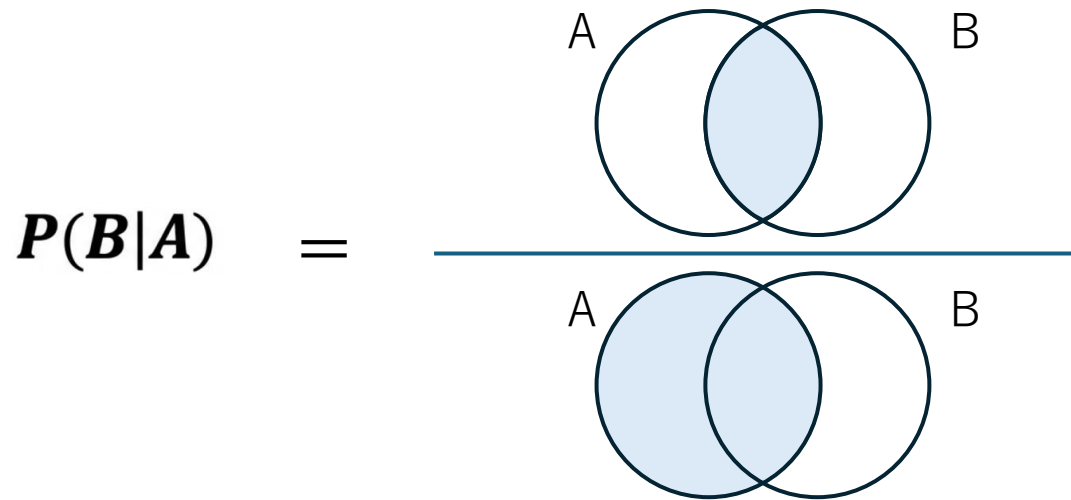
$$\textit{Support}(A, B) = P(A \cap B)$$



信頼度

Aが起きた時、Bが起きる確率

$$\textit{Confidence}(A, B) = P(B|A)$$



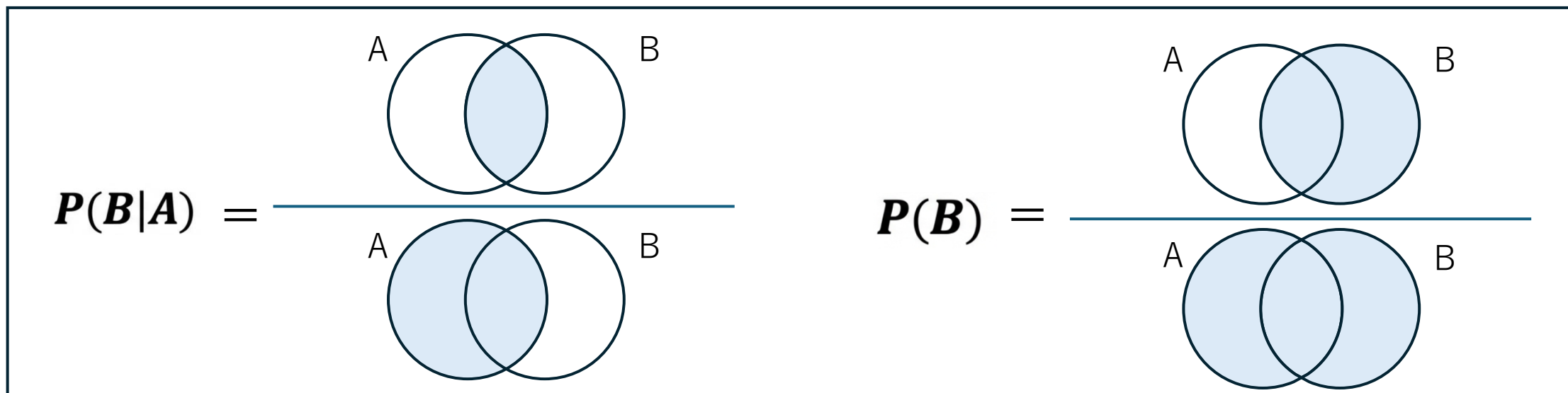
意思決定のカギとなる指標

リフト値

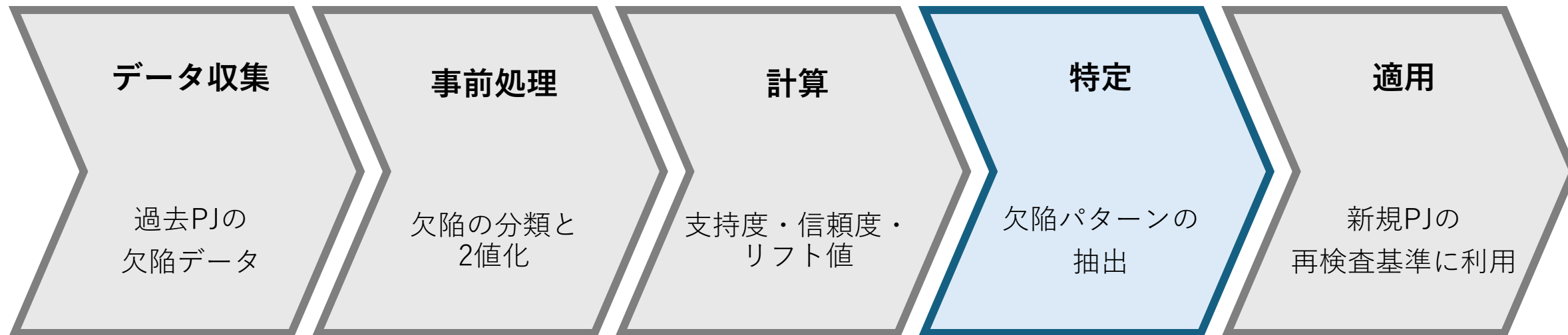
「全体でBが発生する確率」に対して
「Aが発生した条件下でBが発生する確率」が何倍かを示す

$$Lift(A, B) = \frac{P(B|A)}{P(B)}$$

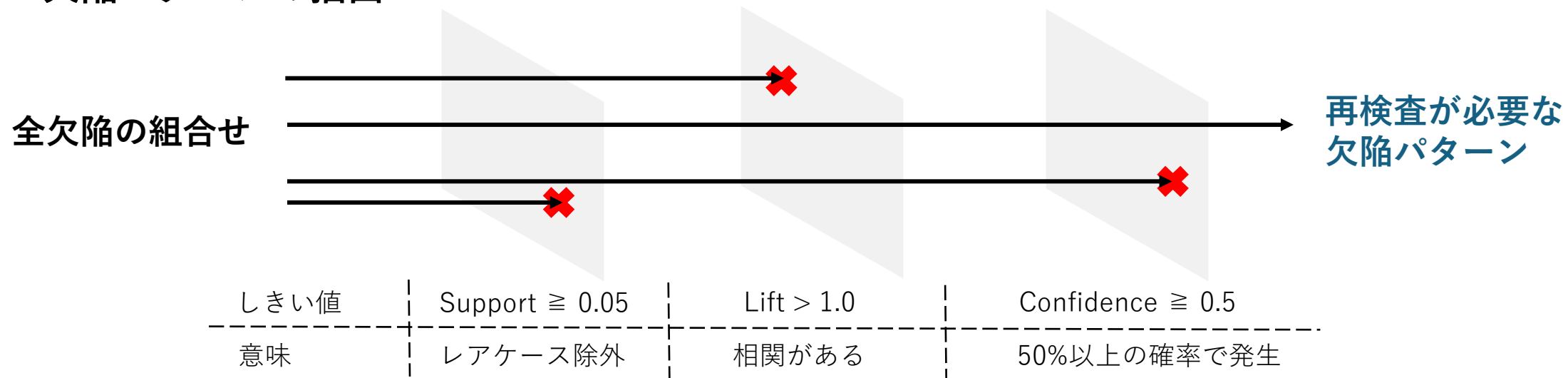
Lift > 1.0 (正の関連) AがあるならBを探索すべき (Action Required)
Lift = 1.0 (関連なし)
Lift < 1.0 (負の関連)



GATEの流れ



欠陥パターンの抽出

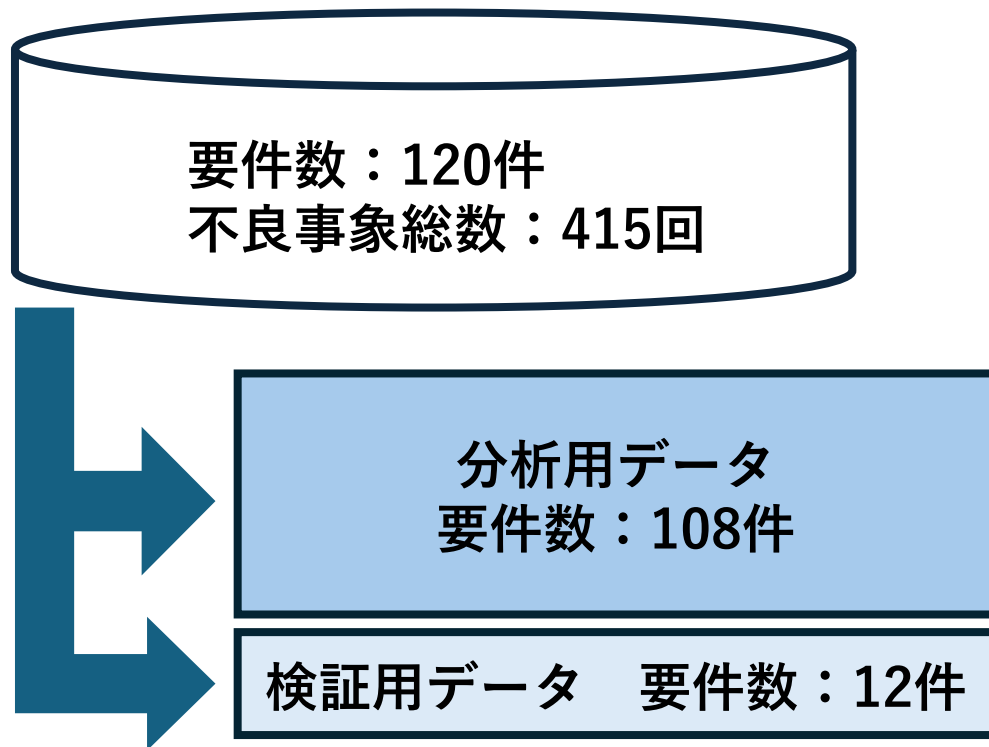


実組織のプロジェクトデータを用いた有効性検証

■実験データ

研究員が所属する組織の実プロジェクトにおけるドキュメント検査の記録を利用

9 : 1 の割合で分析用と検証用のデータ分割して利用



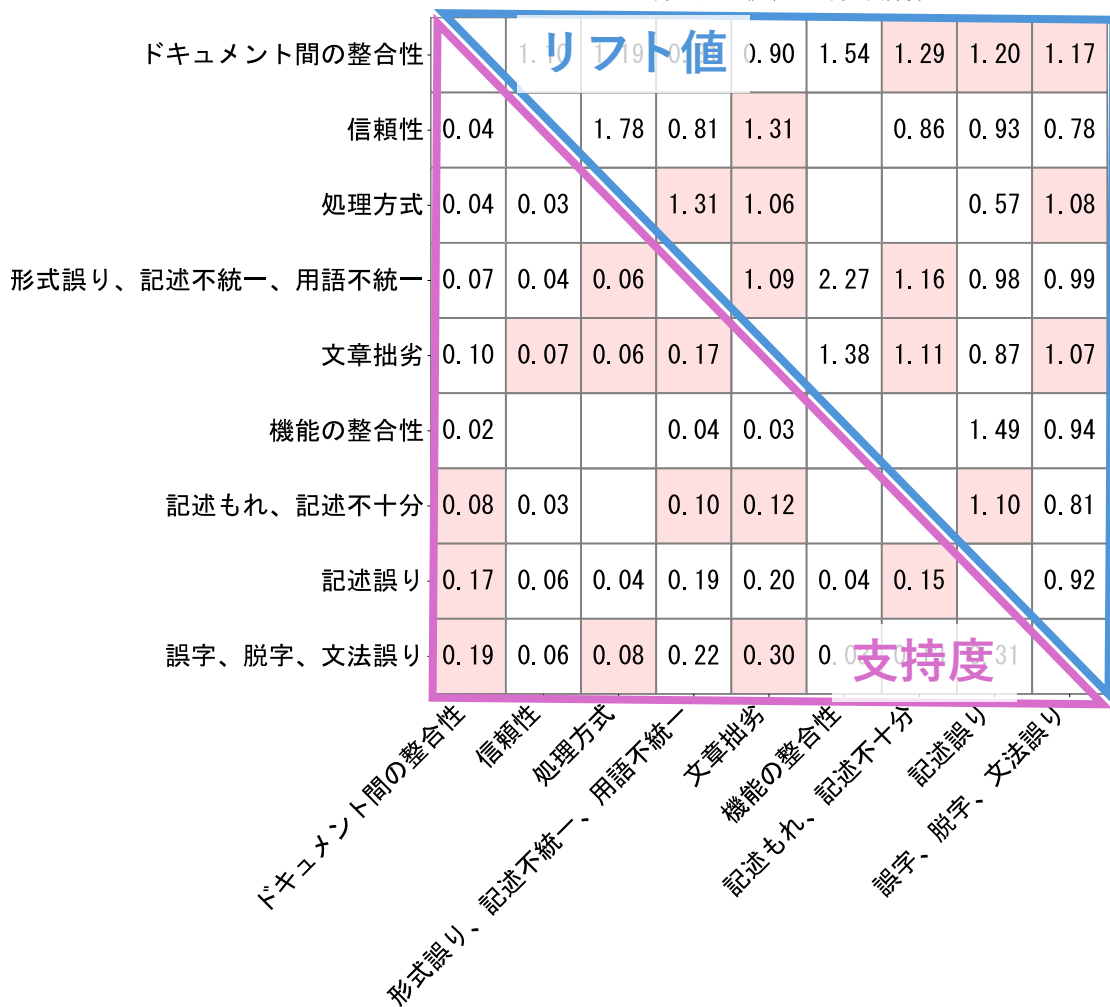
■検査項目

保守性	移植性 記載内容の十分性・妥当性
信頼性	ドキュメント間の整合性
処理方式	誤字、脱字、文法誤り
記述誤り	記述もれ、記述不十分
文章拙劣	形式誤り、記述不統一、用語不統一
	機能の整合性

分析用データによる欠陥パターンの抽出

■欠陥 α があるドキュメントに、欠陥 β もあるか？

欠陥関連性統合マトリックス
上三角：リフト値 / 下三角：支持度



欠陥 α	欠陥 β	$confidence(\alpha, \beta)$
信頼性	文章拙劣	0.57 ✓
ドキュメント間の整合性	記述誤り	0.64 ✓
ドキュメント間の整合性	誤字、脱字、文法誤り	0.75 ✓
記述もれ、記述不十分	記述誤り	0.59 ✓
処理方式	誤字、脱字、文法誤り	0.69 ✓
文章拙劣	誤字、脱字、文法誤り	0.68 ✓
文章拙劣	信頼性	0.17 ✗
記述誤り	ドキュメント間の整合性	0.31 ✗
誤字、脱字、文法誤り	ドキュメント間の整合性	0.30 ✗
記述誤り	記述もれ、記述不十分	0.28 ✗
誤字、脱字、文法誤り	処理方式	0.13 ✗
誤字、脱字、文法誤り	文章拙劣	0.46 ✗

抽出された欠陥パターン

抽出された欠陥パターンの検証

■GATEのしきい値による欠陥パターン特定

	欠陥 α	欠陥 β	$support(\alpha, \beta)$	$lift(\alpha, \beta)$	$confidence(\alpha, \beta)$
分析用	信頼性	文章拙劣	0.07	1.31	<u>0.57</u>
	ドキュメント間の整合性	記述誤り	0.15	1.20	<u>0.64</u>
	ドキュメント間の整合性	誤字, 脱字, 文法誤り	0.18	1.17	<u>0.75</u>
	記述もれ, 記述不十分	記述誤り	0.13	1.10	<u>0.59</u>
	処理方式	誤字, 脱字, 文法誤り	0.08	1.08	<u>0.69</u>
	文章拙劣	誤字, 脱字, 文法誤り	0.27	1.07	<u>0.68</u>

■検証用データの欠陥パターンに対するアソシエーション分析結果

	欠陥 α	欠陥 β	$support(\alpha, \beta)$	$lift(\alpha, \beta)$	$confidence(\alpha, \beta)$
検証用	信頼性	文章拙劣	0.01	<u>4.00</u>	<u>1.00</u>
	ドキュメント間の整合性	記述誤り	0.01	<u>2.00</u>	<u>1.00</u>
	ドキュメント間の整合性	誤字, 脱字, 文法誤り	0.01	<u>1.50</u>	<u>1.00</u>
	記述もれ, 記述不十分	記述誤り	0.04	<u>1.25</u>	<u>0.63</u>
	処理方式	誤字, 脱字, 文法誤り	0.03	<u>1.50</u>	<u>1.00</u>
	文章拙劣	誤字, 脱字, 文法誤り	0.02	1.00	<u>0.67</u>

検証用データにおいても高いリフト値, 信頼度となっていることを確認できた

→欠陥 α があるドキュメントには、欠陥 β が潜んでいる

より精度を上げるためには？

■ GATEのしきい値による欠陥パターン特定

	欠陥 α	欠陥 β	$support(\alpha, \beta)$	$lift(\alpha, \beta)$	$confidence(\alpha, \beta)$
分析用	信頼性	文章拙劣	0.07	1.31	<u>0.57</u>
	ドキュメント間の整合性	記述誤り	0.15	1.20	<u>0.64</u>
	ドキュメント間の整合性	誤字, 脱字, 文法誤り	0.18	1.17	<u>0.75</u>
	記述もれ, 記述不十分	記述誤り	0.13	1.10	<u>0.59</u>
	処理方式			1.08	<u>0.69</u>
	文章拙劣			1.07	<u>0.68</u>

Lift ≥ 1.1 のパターンのみ抽出

■ 検査用データの欠陥パターンに対するアソシエーション分析結果

	欠陥 α	欠陥 β	$support(\alpha, \beta)$	$lift(\alpha, \beta)$	$confidence(\alpha, \beta)$
検証用	信頼性	文章拙劣	0.01	<u>4.00</u>	<u>1.00</u>
	ドキュメント間の整合性	記述誤り	0.01	<u>2.00</u>	<u>1.00</u>
	ドキュメント間の整合性	誤字, 脱字, 文法誤り	0.01	<u>1.50</u>	<u>1.00</u>
	記述もれ, 記述不十分	記述誤り	0.04	<u>1.25</u>	<u>0.63</u>
	処理方式			1.50	<u>1.00</u>
	文章拙劣			<u>1.00</u>	<u>0.67</u>

Lift=1.0のパターンを排除できた

より厳しい閾値を設けることで、関連の弱いパターンを排除しやすくなる

GATEは欠陥パターンの抽出に有効だった

- 組織特有の欠陥パターンを可視化できる
- 客観的な数値で再検査を決定できる



未検出欠陥の流出リスクを軽減できる

GATEをより有効な手法にするために

データ蓄積の自動化

サンプル数を増やし統計的信頼性を向上
導入時の負担を軽減



現場導入

実プロジェクトでの有効性を検証
現場のフィードバックをもとに手法改善



欠陥の要因も考慮した分析

なぜ欠陥パターンが存在するのかを解明
根本から品質改善

GATEで品質基準を明確にしよう！



以上、ありがとうございました