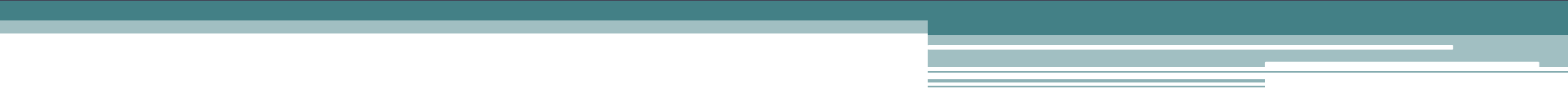


鉄道利用者の所要時間信頼性 向上便益の推計手法の開発

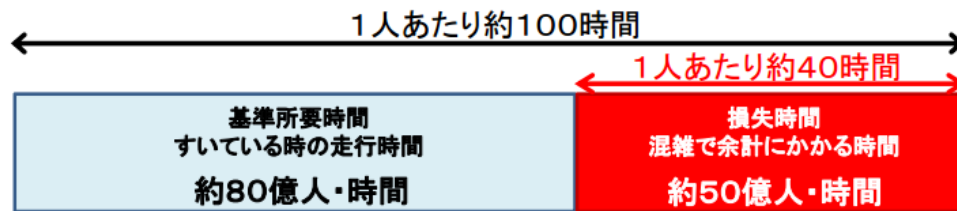
東京電機大学 理工学部
建築・都市環境学系
高田 和幸

研究の背景（動機）



道路の渋滞損失

○ 一人あたりの年間渋滞損失時間は約40時間で、乗車時間(約100時間)の約4割に相当。

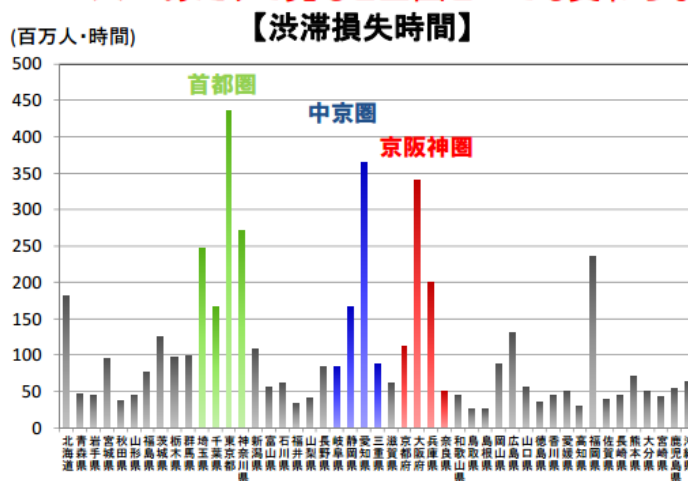


欧米の主要都市における渋滞損失は移動時間の約2割

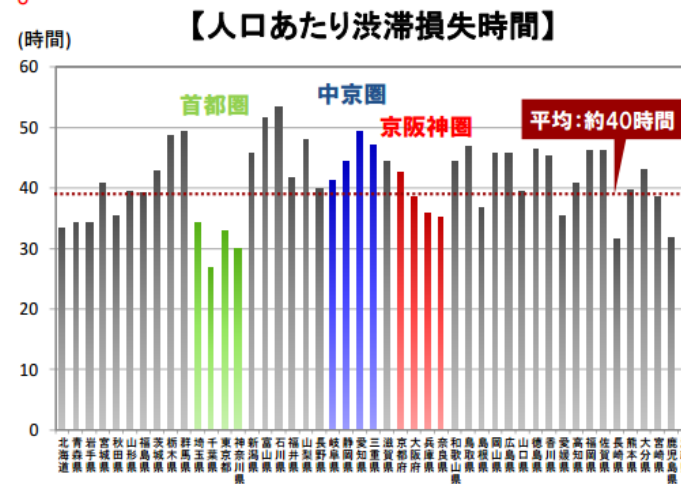
出典) TomTom Americas Traffic Index
TomTom European Traffic Index

出典) 渋滞損失時間はH24年度プローブデータ、人口は総務省統計資料(H24.10) **約4割**

○ 渋滞損失時間は、都道府県別の総量では三大都市圏等の都市部が突出するが、人口あたりで見ると全国どこでも変わらない。

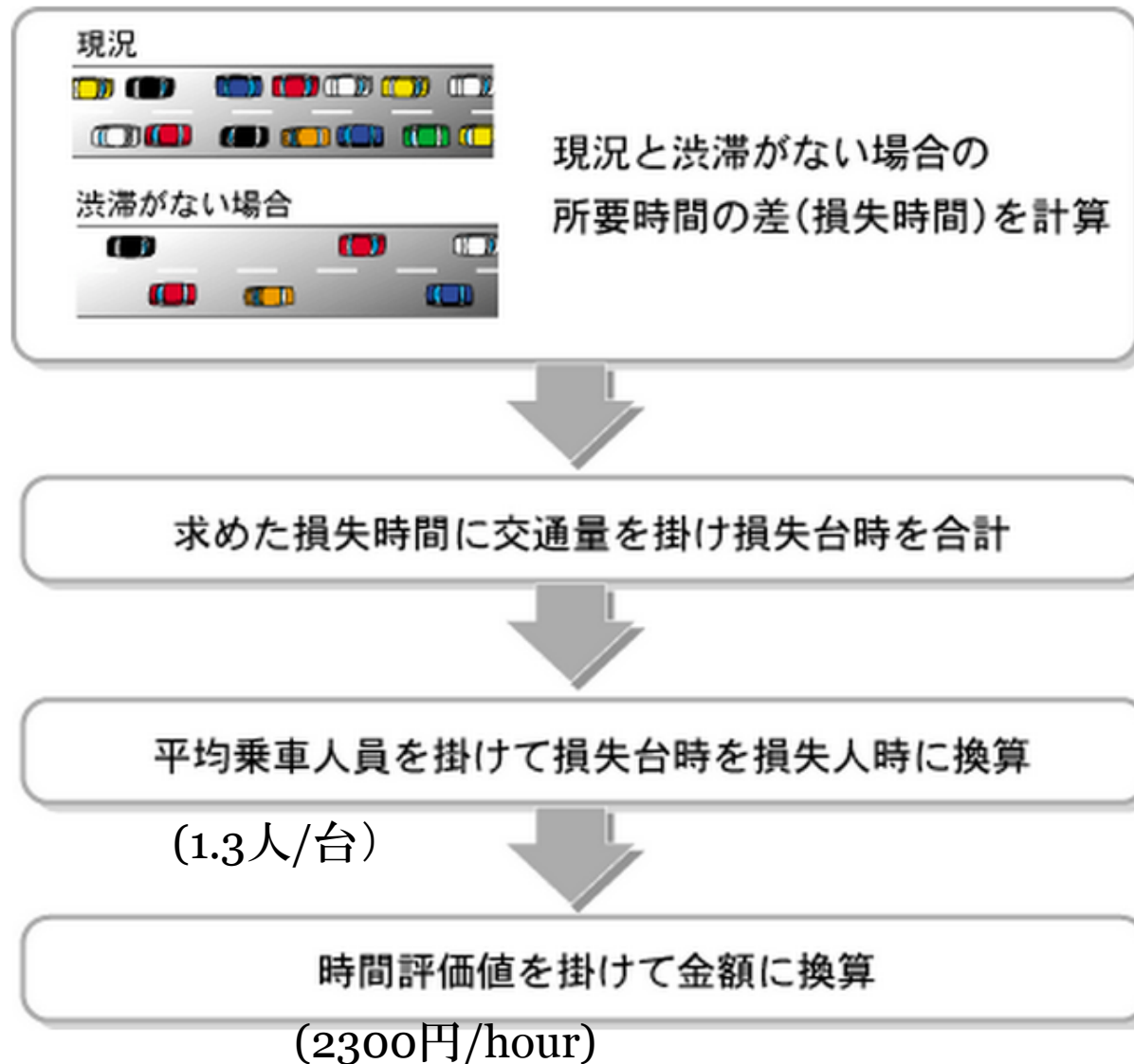


出典) 渋滞損失時間はH24年度プローブデータ、人口は総務省統計資料(H24.10)



ref. 交通流対策について／国土交通省 道路局（平成27年3月5日）

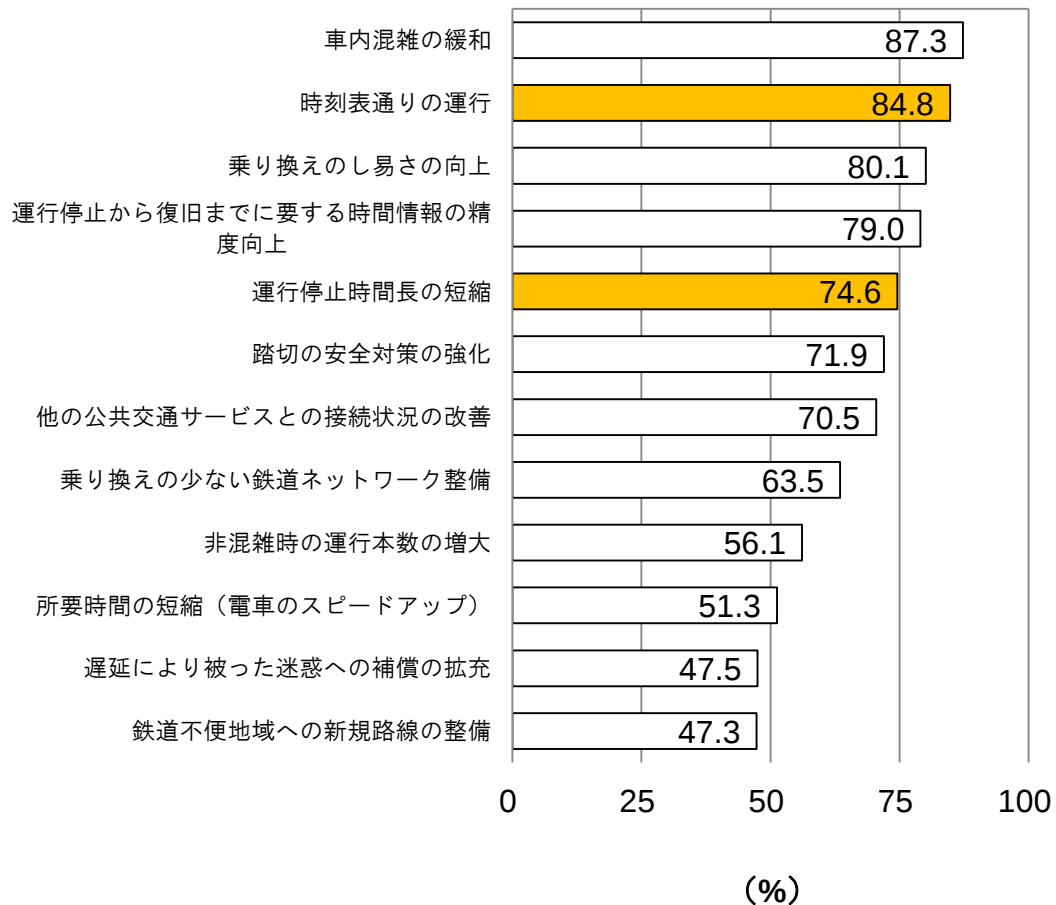
渋滞損失時間、渋滞損失額の求め方のフロー



2006年に発生した主な鉄道輸送障害

月	日	時刻	路線名	原因	影響	影響人数
2	16	18:16	小田急線	人身事故	全線運休	10,000
2	21	10:05	井の頭線	車止め衝突事故	全線運休	
3	11	10:54	大井町線	ドア誤開	全線運休	5,000
4	5	21:30	京王線	踏切事故		
4	6	10:00		車両接触事故		
4	14	17:05	ゆりかもめ	脱輪事故	全線運休	
4	24	10:30	湘南新宿ライン	線路隆起	全線終日運休	320,000
6	13	9:37	都交荒川線	追突事故		
6	21	21:15	千葉都市モノレール2号線	クレーン車接触事故		
6	22	7:19	田園都市線	車両接触事故		
6	22	7:42	東横線	車両故障	全線運休	
8	14	7:37		送電線接触事故	埼京線・中央線・総武線・京葉線 武蔵野線・銀座線・丸の内線・日比谷線 東西線・千代田線・有楽町線・南北線 浅草線・三田線・新宿線・大江戸線 ゆりかもめ・京王線・井の頭線・京成	345,000
9	28	4:20	京葉線	電気系統火災	全線運休 440本運休 (1時間後に西船橋～蘇我運転再開)	220,000
11	24	3:30	東京モノレール	工事用車両故障	全線運休 225本運休	56,500
11	24	21:35	東上線	交通事故	東松山～高坂運休 4本運休	
11	25	0:55	東上線	人身事故	当該列車15分遅れ	
11	27	19:20	武蔵野線	貨物列車停電	全線運休	35,000
11	27	23:20	京王線	踏切事故・脱線事故	全線運休	
11	28	7:10	常磐線	ポイント故障	4本運休11本遅れ	30,000

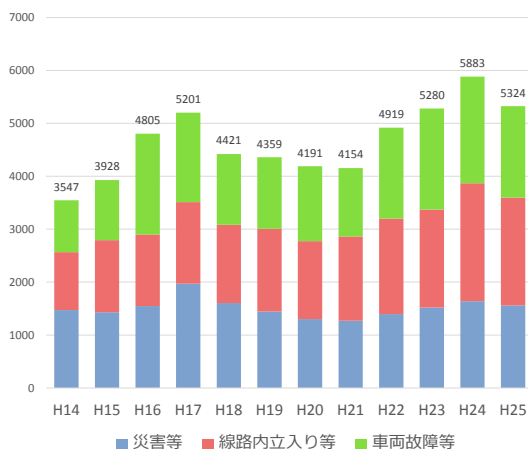
鉄道サービスの改善必要度 (インターネット調査：標本数：1050)



列車遅延・輸送障害対策の評価手法が追記

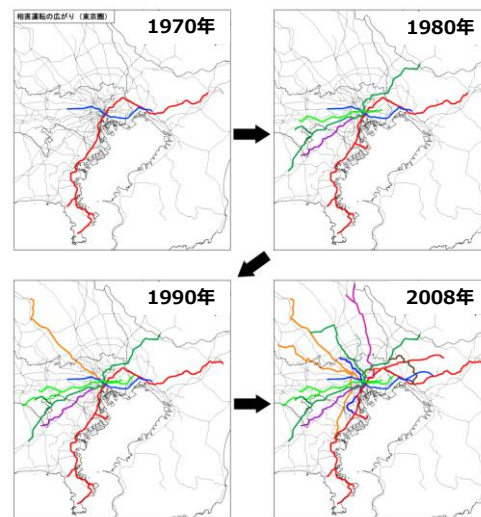
首都圏の鉄道の遅延は増加傾向→所要時間の信頼性が低下

輸送障害件数の推移



国土交通省「輸送障害件数の推移」

相互直通運転路線数の推移



国土交通省「平成20年度 国土交通白書」

輸送障害の発生， 遅延が離れた路線まで波及

• 仮屋崎 圭司，日比野 直彦，森地 茂「都市鉄道の列車遅延の拡大メカニズムに関する研究」『土木計画学研究・論文集』第27巻，第4号，871～880頁，2010年9月

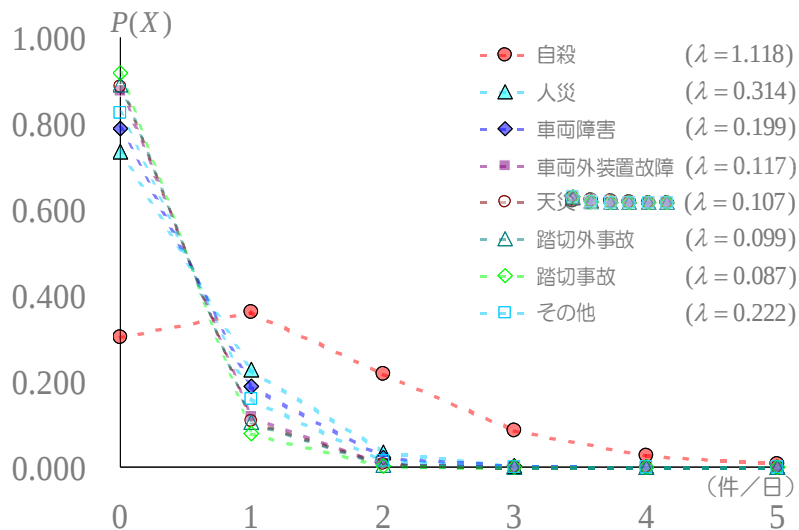
遅延の慢性化 所要時間の信頼性低下

所要時間の信頼性を 向上させる取り組みが必要

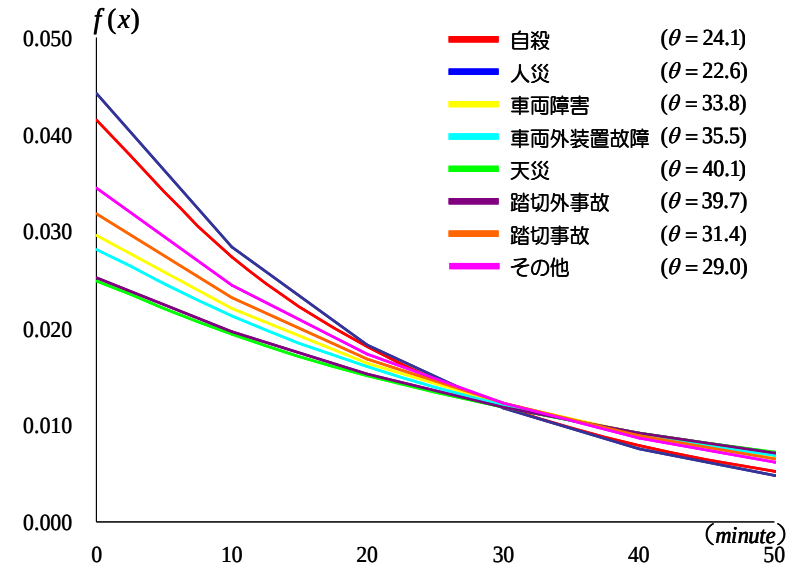
輸送障害に伴う時間損失の推計

- 高田和幸, 小林繭美: 鉄道輸送障害発生時の乗客の選択行動に関する分析, 土木計画学研究・論文集, 25(3), 763-768, 2008.
- 高田和幸, 吉澤智幸: [鉄道事故に伴う旅客の損失時間の推計手法に関する研究](#), 土木計画学研究・論文集, 22, 863-868, 2005.
- Kazuyuki TAKADA, Tomoyuki YOSHIZAWA, Makoto FUJIU: Reliability Evaluation of the Railway Network in the Tokyo Metropolitan Area, 2nd International Symposium on Transport Network Reliability, pp.104 -107, 2004.

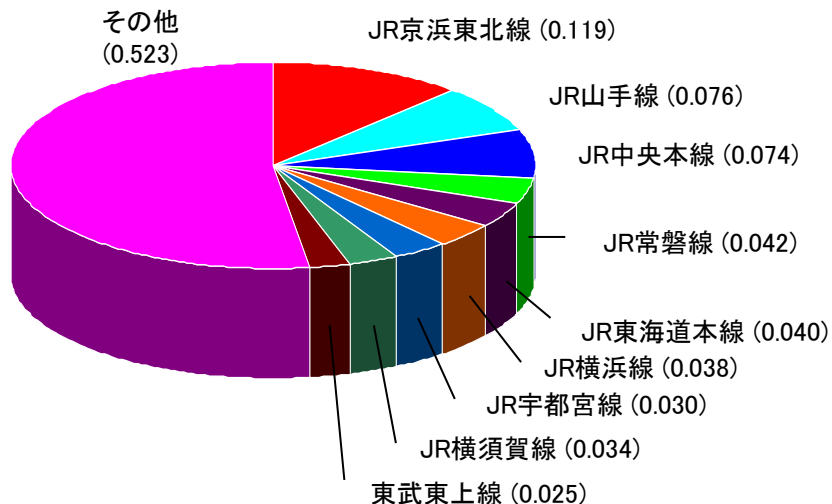
事故発生件数の確率分布



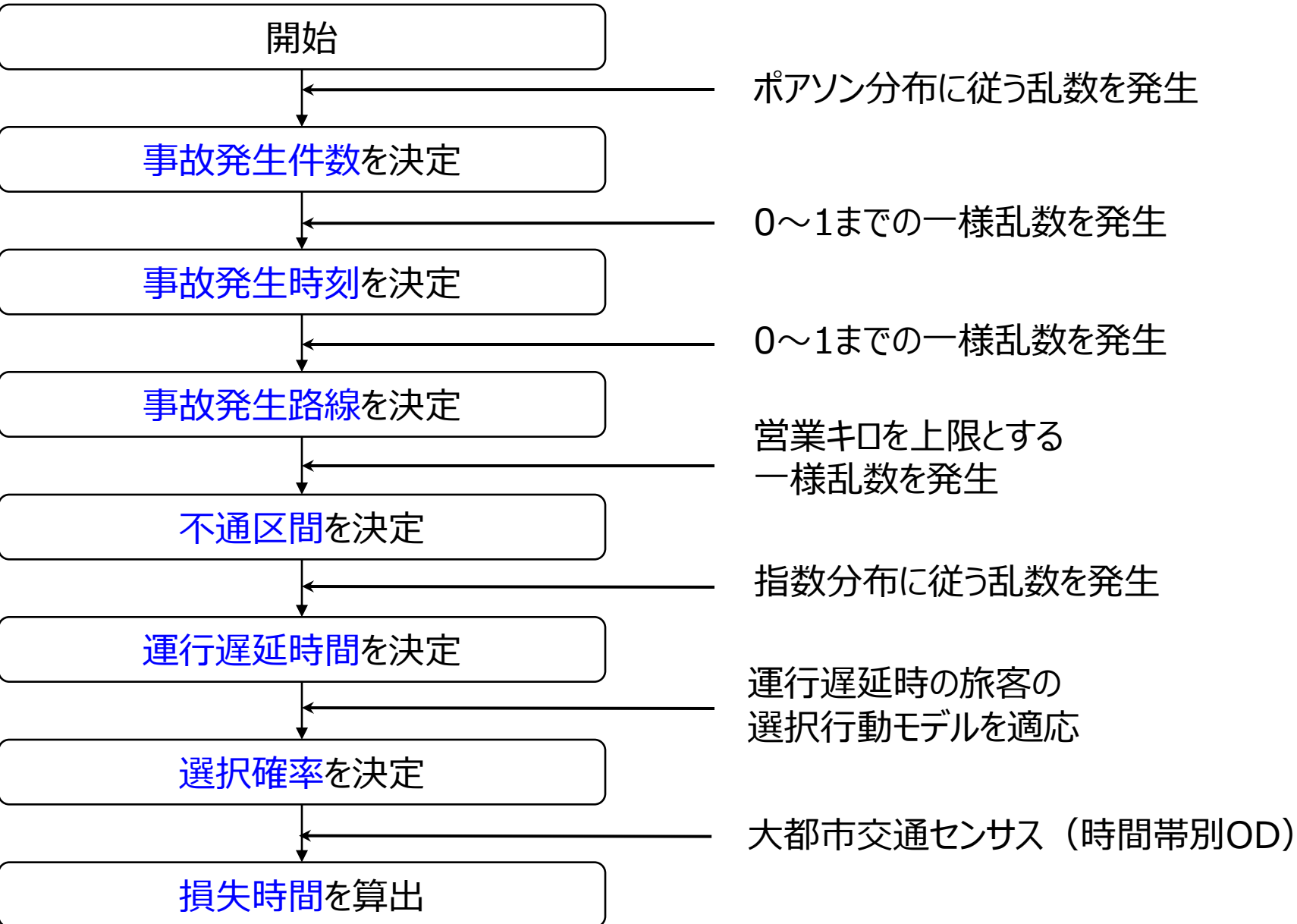
運行遅延時間の分布



事故発生路線の割合



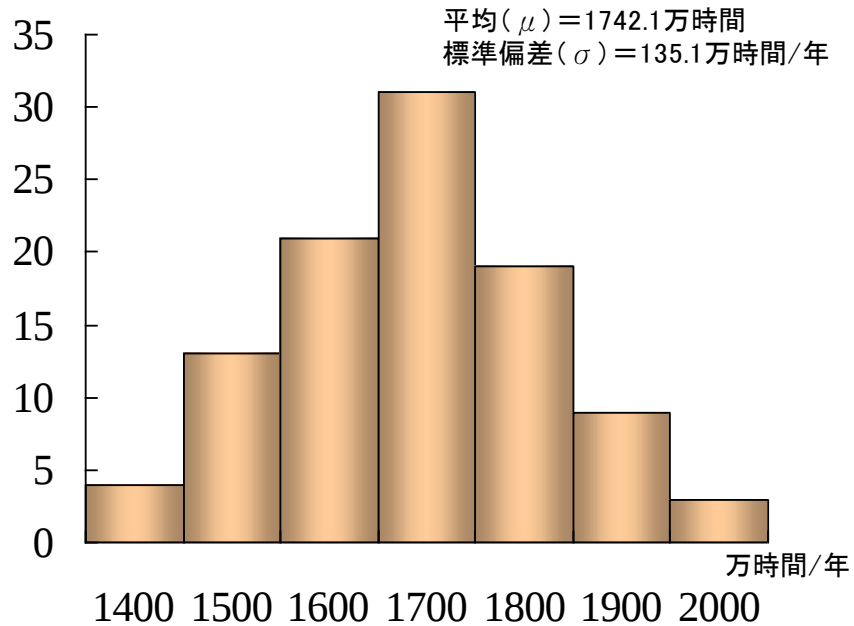
損失時間（到着遅延時間）を推計するシミュレーションのフロー



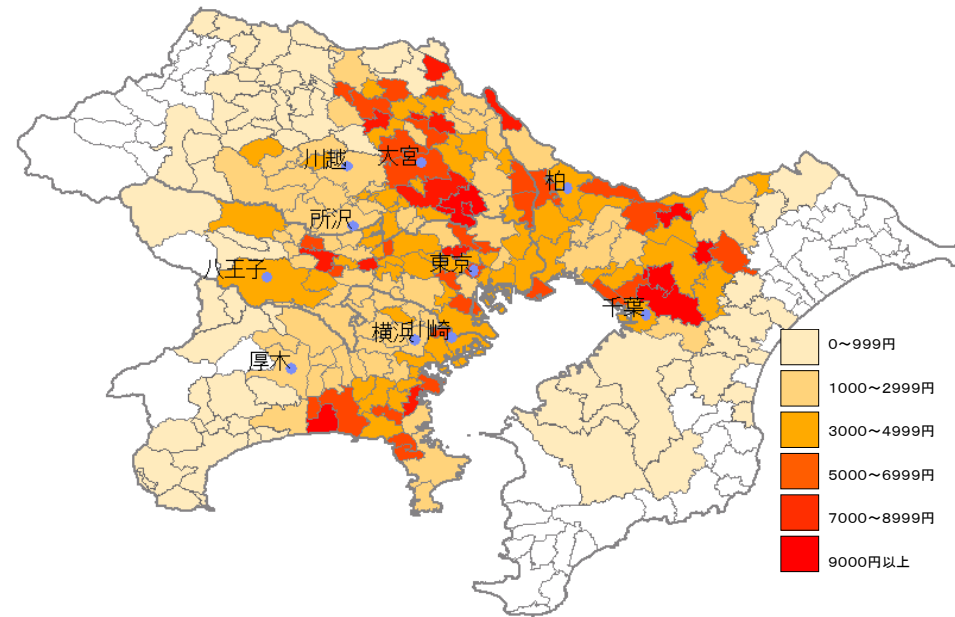
損失時間の推計結果

首都圏全体：**年間1742.1万時間**損失

旅客1人あたり：**年間約2時間**損失



1人あたりの時間損失額の地域分布



時間損失額：400.66 億円/年（時間評価値2300円）

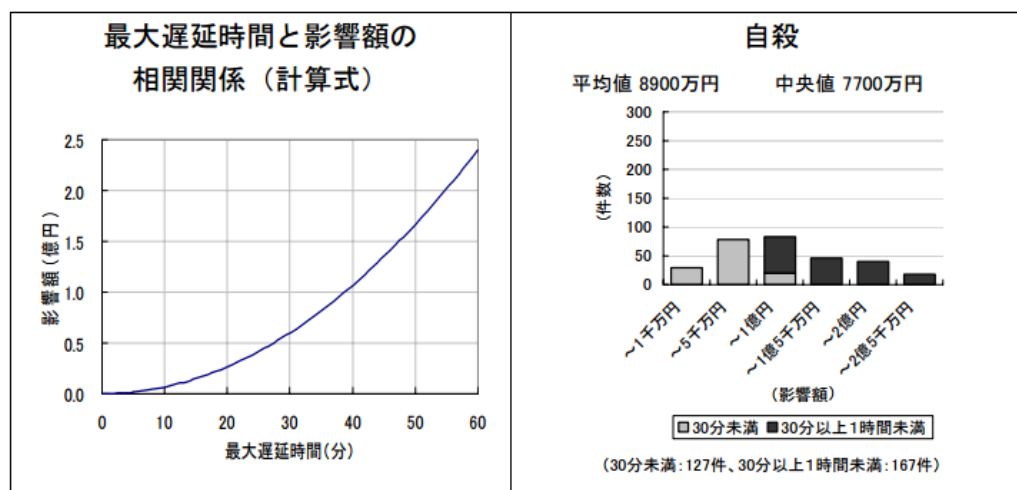
国土交通省鉄道局

鉄道輸送トラブルによる影響に関する調査結果の概要

一大都市圏の1時間未満の輸送トラブルについて(平成21年8月28日)

(座長：屋井 鉄雄 東京工業大学大学院教授)

図5 原因別影響額の概略試算(1件当たり、首都圏、平成19年度)



注1) 鉄道運転事故は含まない。

注2) 対象：首都圏15事業者22路線

注3) 支障時間が1時間未満の輸送トラブルを対象とした。

注4) 影響額： k (総遅延時間概算係数) $\times$ 最大遅延時間² $\times$ 列車当たり輸送人員 $\times$ 時間価値 $\times$ 平均乗車キロ割合

注5) k は混雑率や運行頻度、ダイヤの回復余力等の路線の特性によって幅があるが、 $k=2.5$ (混雑率180%程度・時間当たり本数20-30本程度の線区で、遅延回復まで2-3時間程度要した事象から算出)として試算。

注6) 列車当たりの輸送人員=2000人(1車両定員130~150人、1列車6~10両、混雑率150~180%を想定)、時間価値=51.3円/分、平均乗車キロ割合(路線別輸送人キロ÷路線別輸送人員÷路線長)=0.26として試算。(遅延列車に乗降する乗客への影響その他要因を考慮し、片線側の列車遅延の影響から算出。)

注7) 図中の平均値は1件当たり影響額の平均、中央値は全件数の影響額を大きさの順に並べた際の中央の値。

支払意志額の特性分析

•Kazuyuki Takada, Makoto Fujiu: [Study of Willingness to Pay for Reducing Lost Time of Railway Users](#), Journal of Eastern Asia Society for Transportation Studies, Vol. 8, 1464-1474, 2010.

支払意志額の調査票の設計

- 鉄道利用者が被っている損失時間の設定： 2 時間, 6 時間, 12 時間
- 損失時間の縮減の設定： 50% 解消, 100% 解消

計 6 パターンの質問を作成

Q6: 鉄道の遅延を改良するには、ホームドアの設置や踏切解消などの整備が必要です。仮に鉄道遅延が完全に解消されるならば、これらの遅延を解消する整備のための基金に年間3000円を協力することができますか。

1. 協力できる(年間3000円)
2. 協力できない(年間3000円)

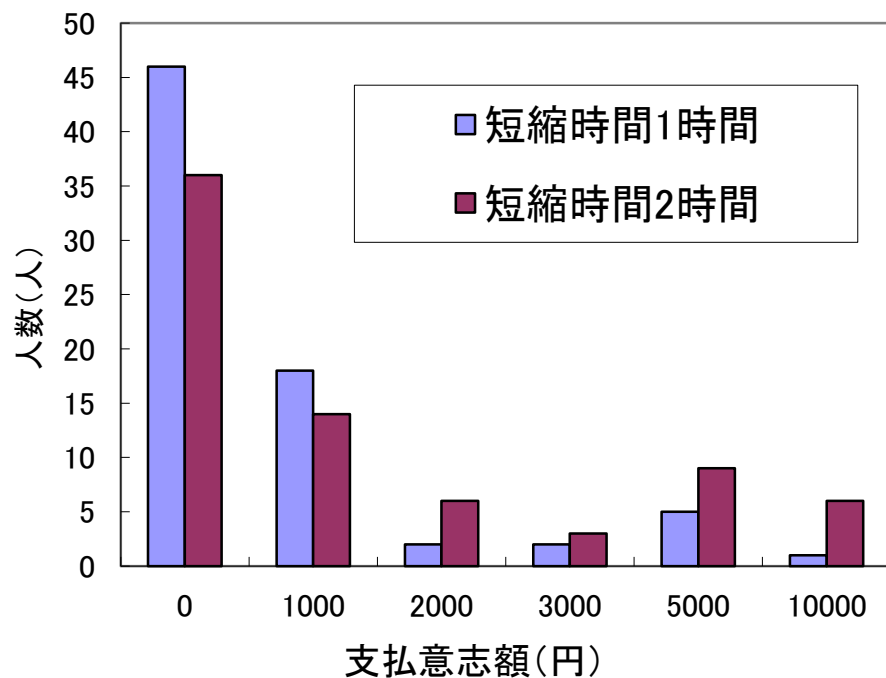
Q7: Q6で回答されたことを踏まえてお答えください。

それでは、遅延を解消する整備のために基金を年間いくらまで協力することができますか。

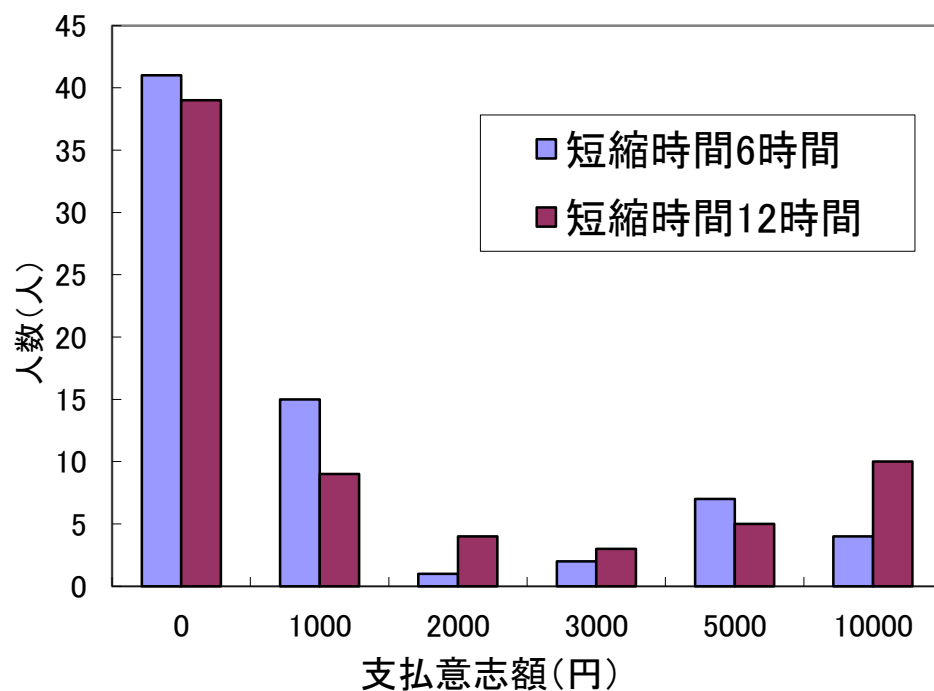
年間 円まで協力することができます。

支払意志額の回答結果

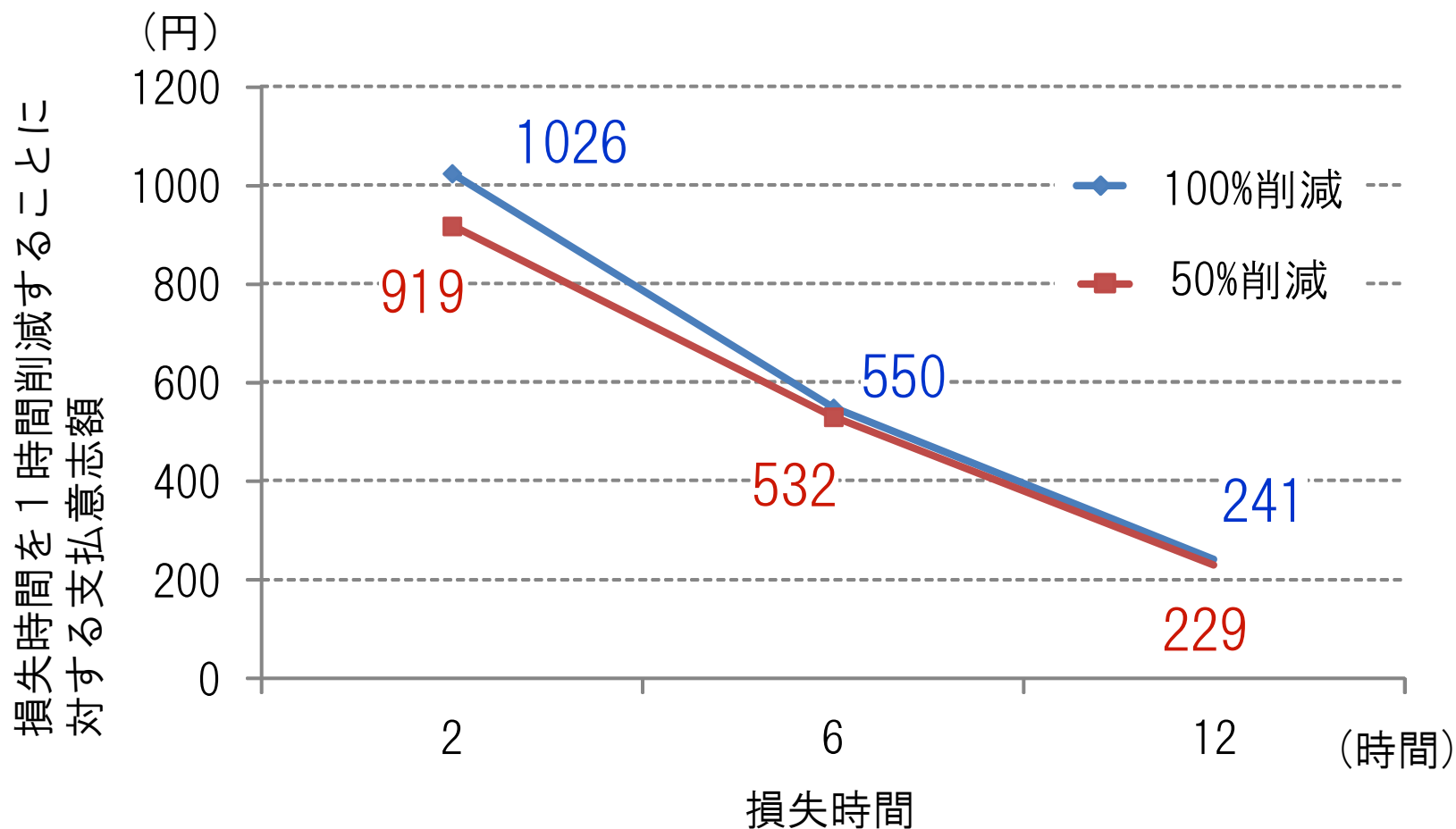
遅延による損失時間が2時間の場合



遅延による損失時間が12時間の場合

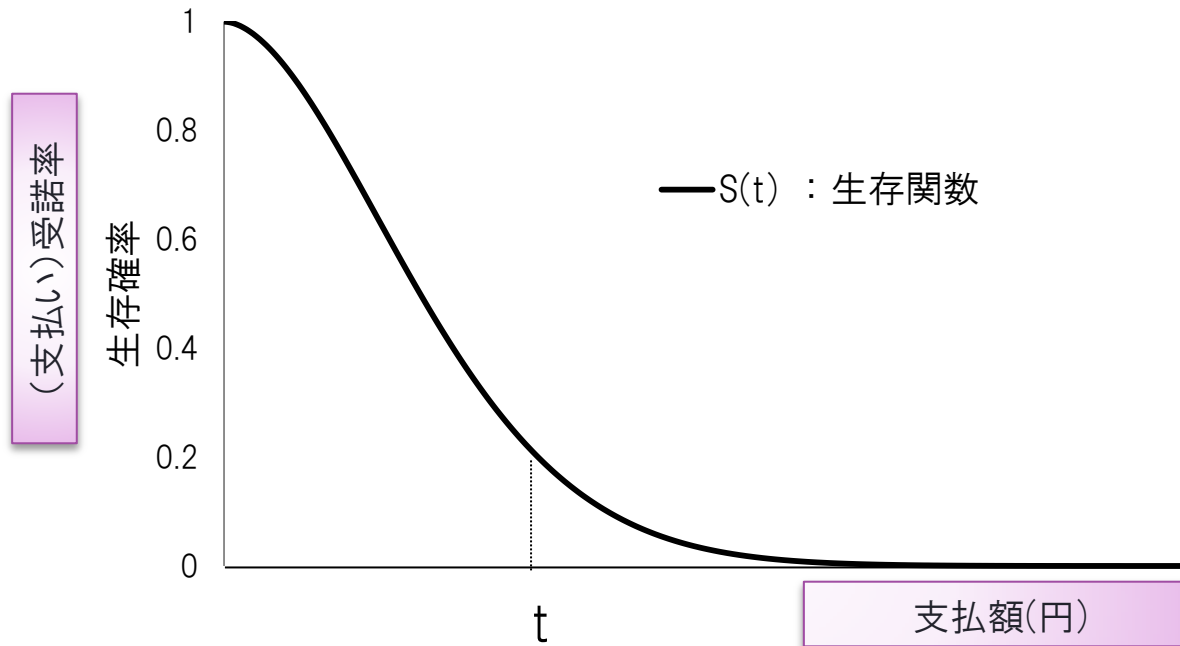


損失時間の削減に対する支払意志額



支払意志額モデルの構築

生存分析（比例ハザードモデル）



支払意志額にワイブル分布を仮定

ハザード関数 :
$$h(t) = \gamma \lambda t^{\gamma-1} \exp\left(\sum_i \beta_i x_i\right)$$

生存関数 :
$$S(t) = \exp\left\{(-\lambda t^\gamma) \exp\left(\sum_i \beta_i x_i\right)\right\}$$

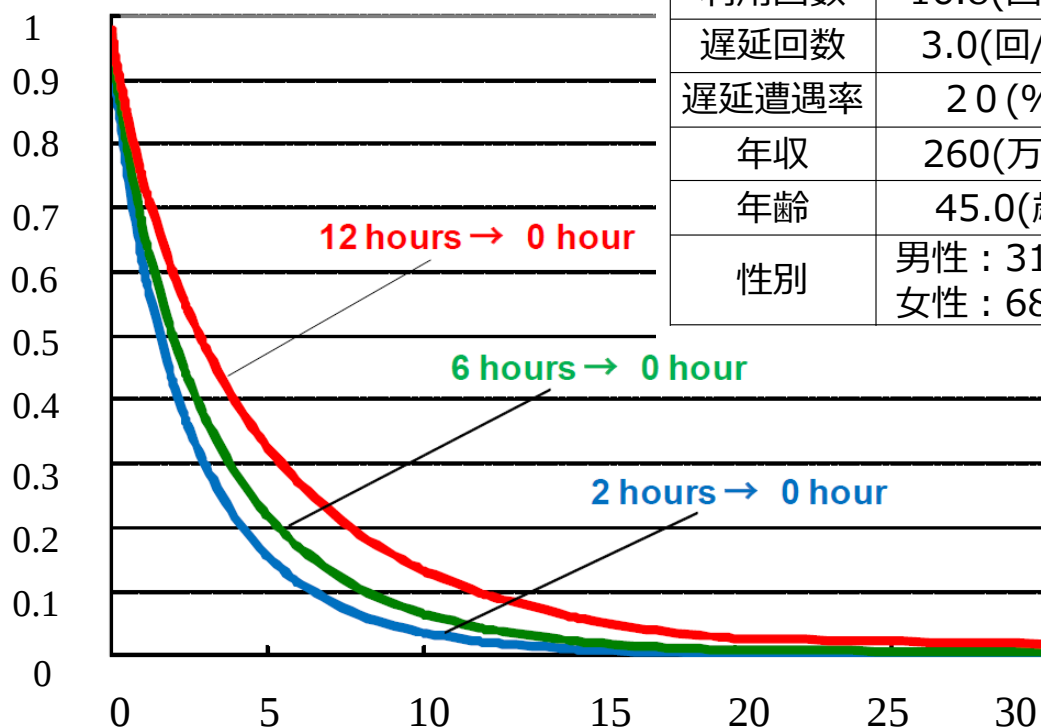
t	: 支払意志額
X	: 共変量 (説明変数)
β	: 共変量のパラメータ
λ	: 形状パラメータ
γ	: スケールパラメータ

ハザード関数の推定結果

生存関数

共変量	単位	係数	t 値
利用回数	回/月	-0.043	-3.91
遅延回数	回/月	0.028	0.85
損失時間	時間	0.061	2.10
縮減時間	時間	-0.11	-3.24
遅延遭遇率	—	-0.878	-3.00
年収	百万円	0.014	0.52
年齢	歳	-0.116	-2.52
性別	—	0.252	1.58
スケールパラメータ	—	0.859	17.53
形状メータ	—	0.0017	16.25
AIC	—	4195	

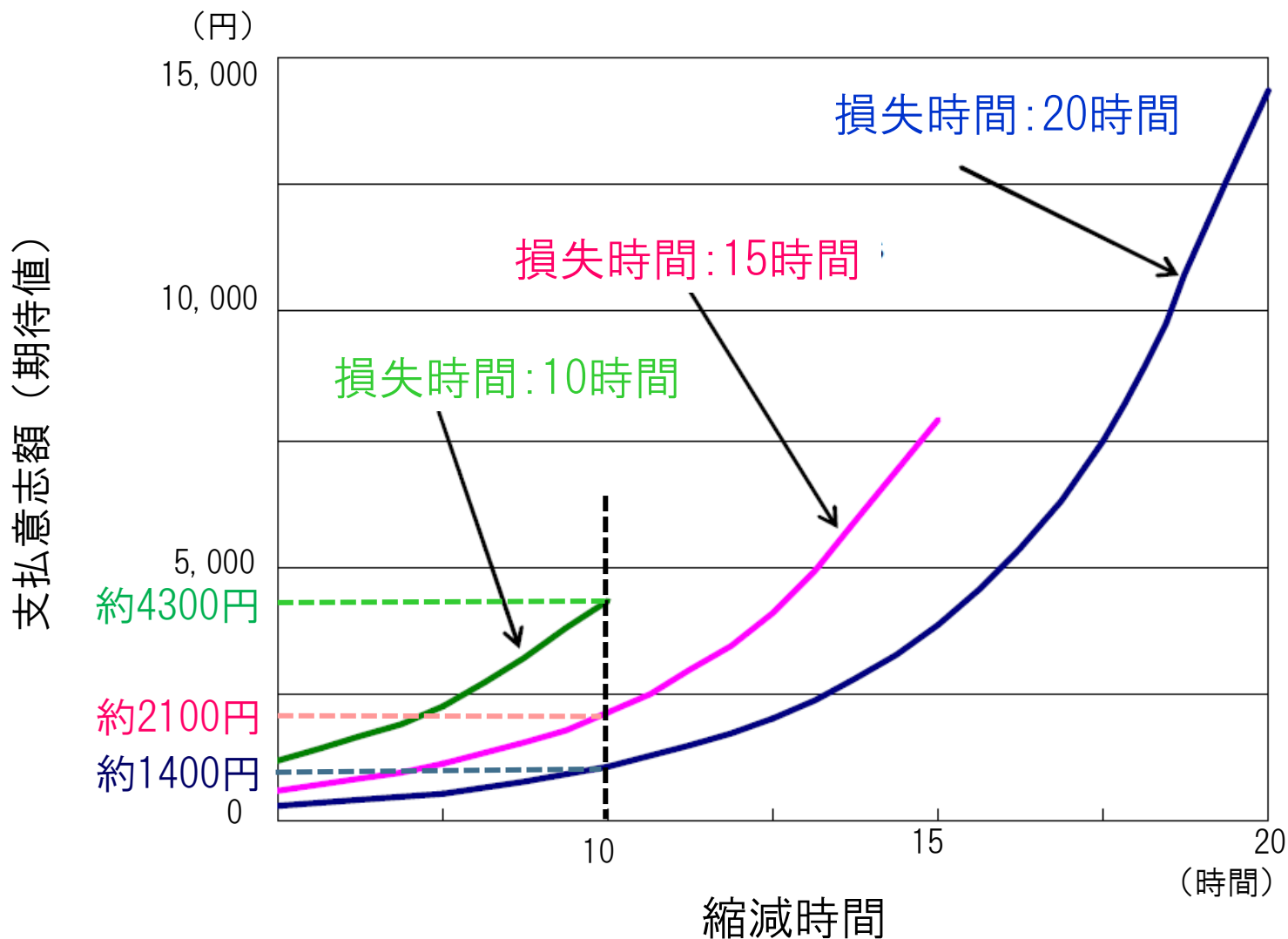
受諾率(×100%)



項目	平均
利用回数	16.8(回/月)
遅延回数	3.0(回/月)
遅延遭遇率	20(%)
年収	260(万円)
年齢	45.0(歳)
性別	男性 : 31.3% 女性 : 68.7%

支払い額
(千円)

損失時間×縮減時間×支払意志額



出発時刻（余裕時間）の決定行動

- Kazuyuki TAKADA, Yuzo TAKANAMI: Estimation of Buffer Time of Rail Commuters in Tokyo Metropolitan Area, EASTS (投稿中)
- 高田和幸, 藤生慎, 横山茂樹, 鈴木孝典: 鉄道の遅延時間を考慮した出発時刻決定行動に関するモデル分析, 土木学会論文集D3 (土木計画学), 68, 5 (土木計画学研究・論文集, 29), 2012.

アンケート調査概要

調査方法	インターネット調査（マクロミル）	
調査年月日	2012年3月28, 29日	
標本数	回収数：1414, （回答数：1221）	
対象者	首都圏の1都3県に居住する鉄道利用者 15歳以上の有職者(パート・アルバイトを含む)	
調査項目	個人属性	性別, 年齢, 未婚既婚, 子供の有無, 居住地, 職業, 年収
	鉄道の利用状況	通勤時の起終点駅(路線名・駅名) 通勤時の経路(乗り換え駅) 鉄道利用頻度 認知している所要時間 遅着・早着の状況
	鉄道遅延経験	運行遅延に関する認識（遅延遭遇頻度, 損失時間）
	支払意志額	鉄道遅延の改善に対する支払意志額
	SP実験	鉄道サービスに関する選好意識（運賃, 所要時間とその変動）

余裕時間の算出過程

アンケート調査

起終点
(高坂-北千住)

乗り換え駅
(池袋, 西日暮里)

出発時刻
(7:25)

希望到着時刻
(終着駅)
(9:00)

	□ 高坂	[駅情報]
07:25~08:19	13駅	東武東上線急行
	○ 池袋	[出口案内] [駅情報]
08:25~08:35	5駅	JR山手線外回り・上野・東京方面
	○ 西日暮里	[出口案内] [駅情報]
08:41~08:47	2駅	東京メトロ千代田線・我孫子行
	□ 北千住	[出口案内] [駅情報]

時刻表(Yahoo!経路)

最短所要時間
(1h22m)

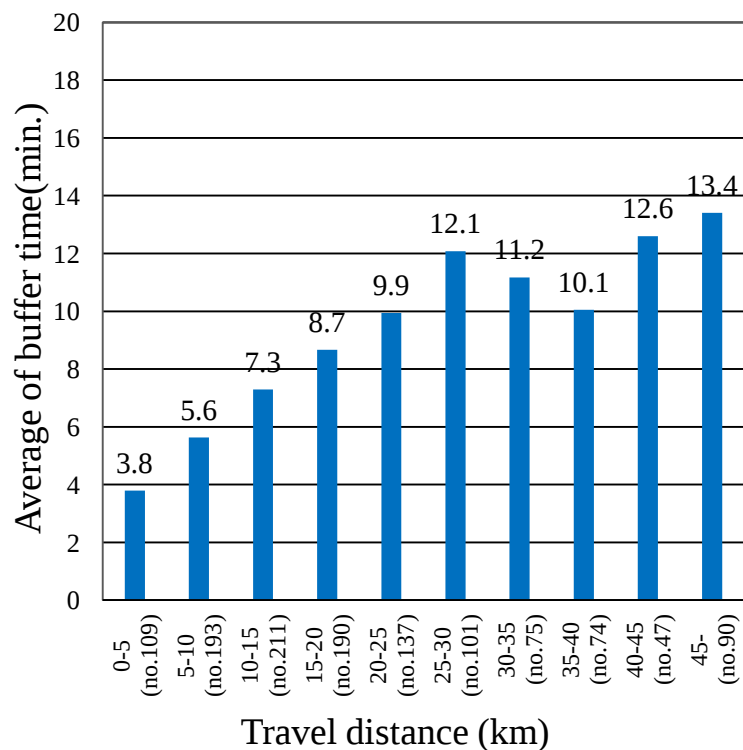
見積もり所要時間
(1h35m)

余裕時間

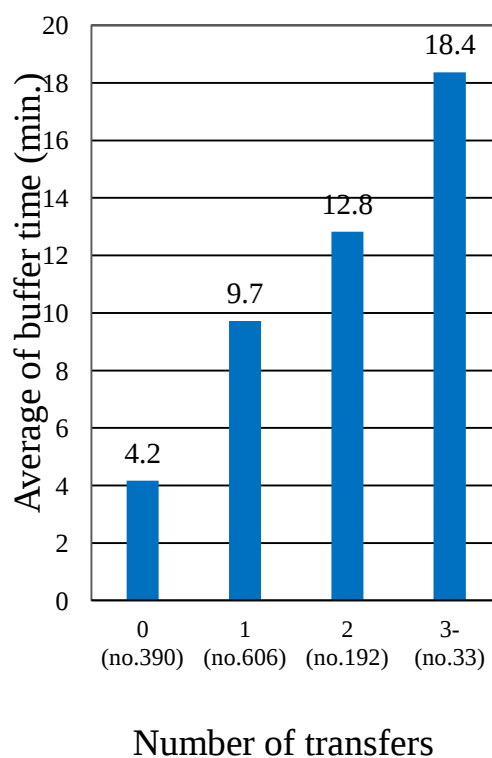
余裕時間
(13 min.)

余裕時間の影響要因

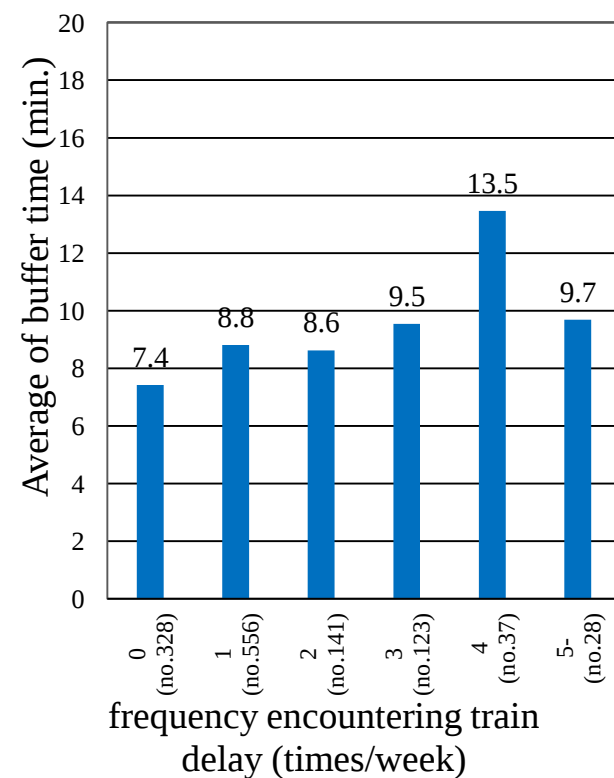
移動距離と余裕時間の関係



乗換回数と余裕時間の関係

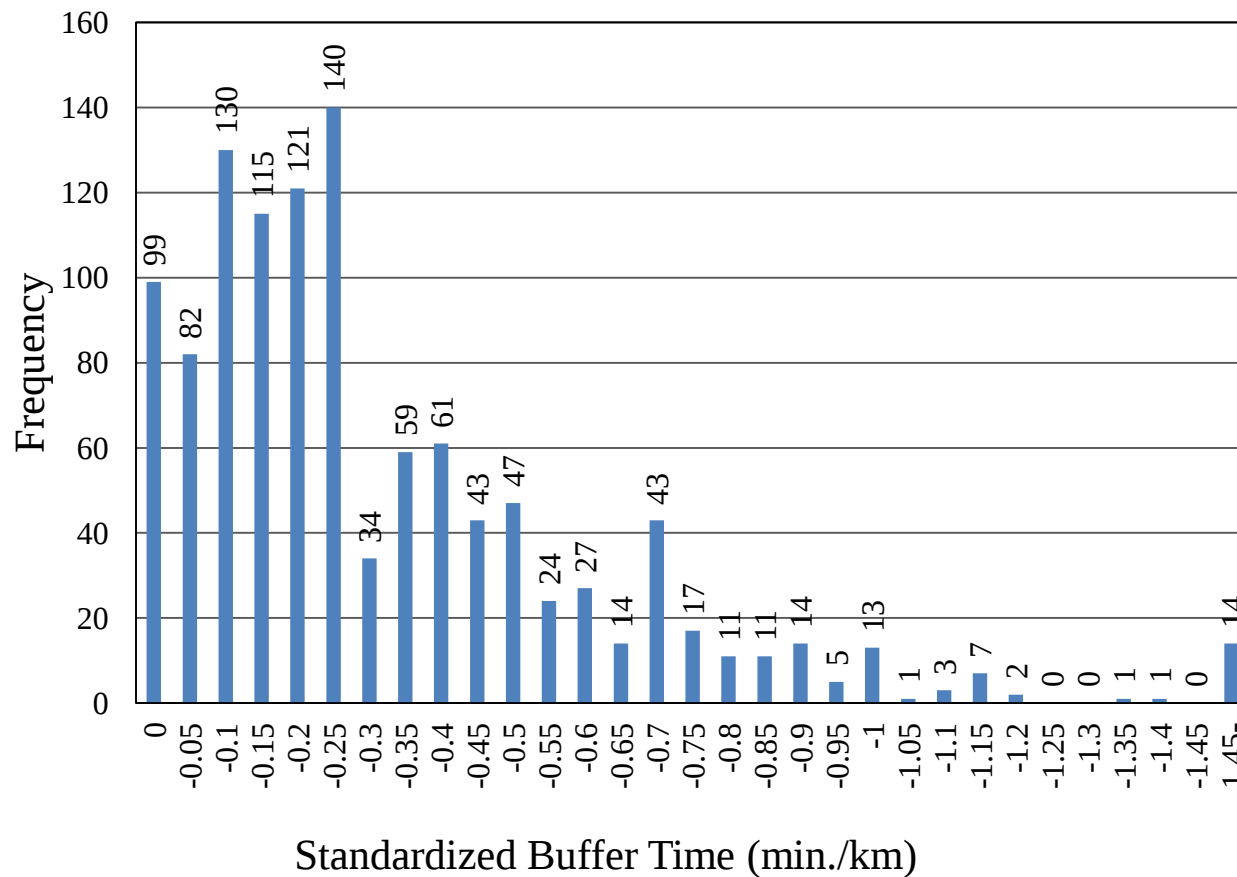


遅延遭遇頻度と余裕時間の関係



余裕時間がマイナスとなった193の標本は除いた

乗車1kmあたり余裕時間



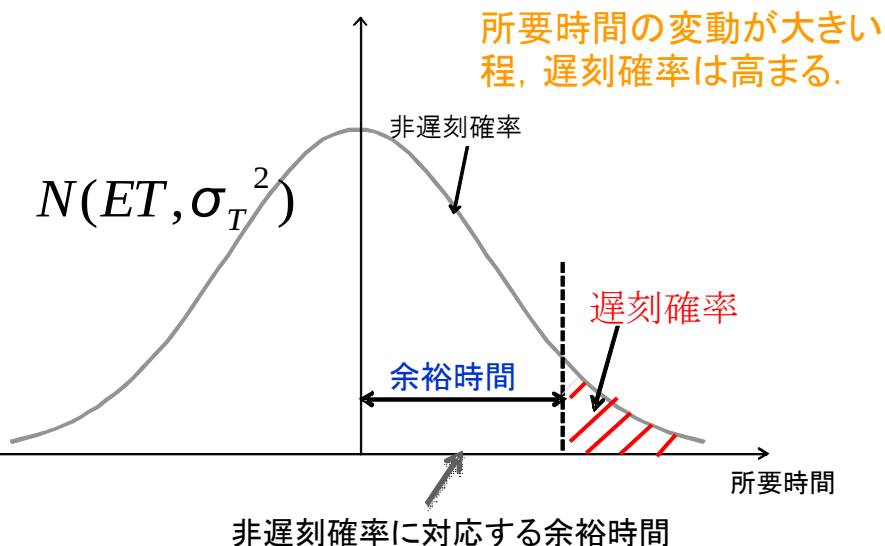
出発時刻決定モデルの構築 (1)

Fosgerauら (2007) の統合アプローチモデルを援用

平均一分散アプローチ

$$U = \delta \cdot C + \alpha \cdot ET + \rho \cdot \sigma_T$$

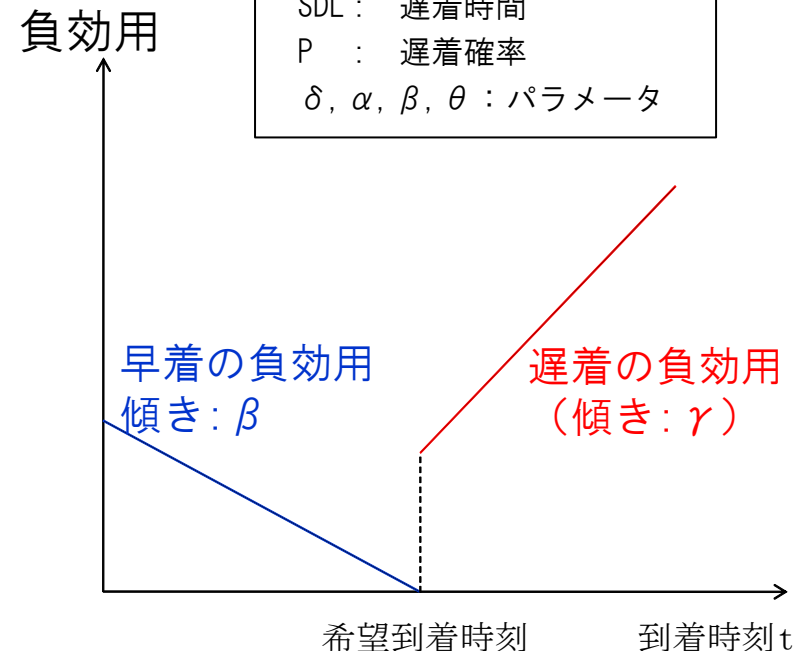
C : 費用
ET : 期待所要時間
 σ_T : 標準偏差
 δ, α, ρ : パラメータ



スケジューリングアプローチ

$$U(t) = \delta \cdot C + \alpha \cdot T + \beta \cdot SDE + \gamma \cdot SDL + \theta \cdot P$$

C : 費用
T : 所要時間
SDE : 早着時間
SDL : 遅着時間
P : 遅着確率
 $\delta, \alpha, \beta, \theta$: パラメータ



出発時刻決定モデルの構築 (2)

統合アプローチモデル

$$U(D, T) = \alpha D + \omega T + \beta(T - D) \quad (1)$$

早発の不効用 所要時間の不効用 遅着の不効用

T : 所要時間
D : 出発時間

所要時間(T)を，時刻表に基づく最短所要時間(μ)と基準化余裕時間(T_s)と移動距離(L)で表現

$$T = \mu + L \cdot T_s \quad (2)$$

基準化余裕時間は指数分布に従い，また複数の因子（共変量）の影響を受けると仮定．

$$h(t|x_i) = \lambda \exp\left(\sum_{i=1}^n \theta_i x_i\right) \quad (3)$$

x_i : covariates, θ_i : parameters, λ : scale parameter.

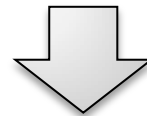
出発時刻決定モデルの構築 (3)

余裕時間の密度関数

$$f(T_s) = \lambda \exp\left(\sum_{i=1}^n \theta_i x_i\right) \cdot \exp\left\{(-\lambda L T_s) \exp\left(\sum_{i=1}^n \theta_i x_i\right)\right\}$$

期待負効用の最小化問題

$$E[U(D)] = \min \left[\alpha D + \omega \mu + \beta \int_{D-\mu}^{\infty} (\mu + L T_s - D) f(T_s) dT_s \right]$$



最適出発時間(見積みり所要時間)

$$D^* = \mu + L F^{-1}\left(1 - \frac{\alpha}{\beta}\right)$$

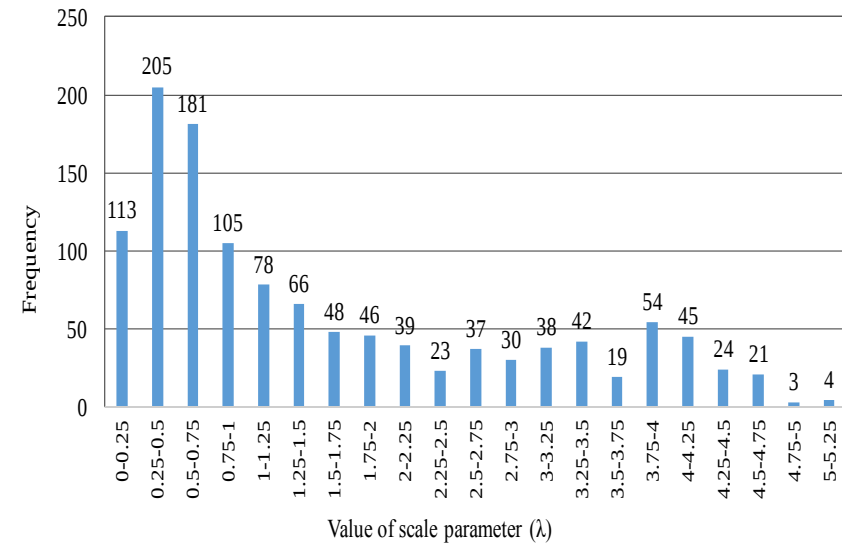
基準化余裕時間

$$\frac{D^* - \mu}{L} = \frac{\ln \frac{\alpha}{\beta}}{\lambda \cdot \exp\left(\sum_{i=1}^n \theta_i x_i\right)} \quad \longrightarrow \quad \text{MCMCで推定}$$

出発時刻決定（余裕時間）モデルの推定結果

λ の分布

	事前分布	パラメータ	t値
最適遅着確率	$\frac{\alpha}{\beta} \sim \text{unif}(0,1)$	0.774**	69.79
乗り換え回数	$\theta_1 \sim \text{unif}(-5,5)$	-0.208**	-5.06
遅延遭遇頻度	$\theta_2 \sim \text{unif}(-5,5)$	-0.0277	-1.73
指数分布パラメータ	$\lambda_i \sim \ln N(0, \sigma^2)$		
λ の分散	$\sigma^2 \sim \text{unif}(0,10)$	1.038**	38.45
基準化余裕時間の分散	$\left(\frac{D^* - \mu}{L}\right) \sim N(0, \delta^2)$ $\delta^2 \sim \text{unif}(0,10)$	0.0740**	10.78
サンプル数	1212		



** : Significant at 1% level; * : Significant at 5% level.

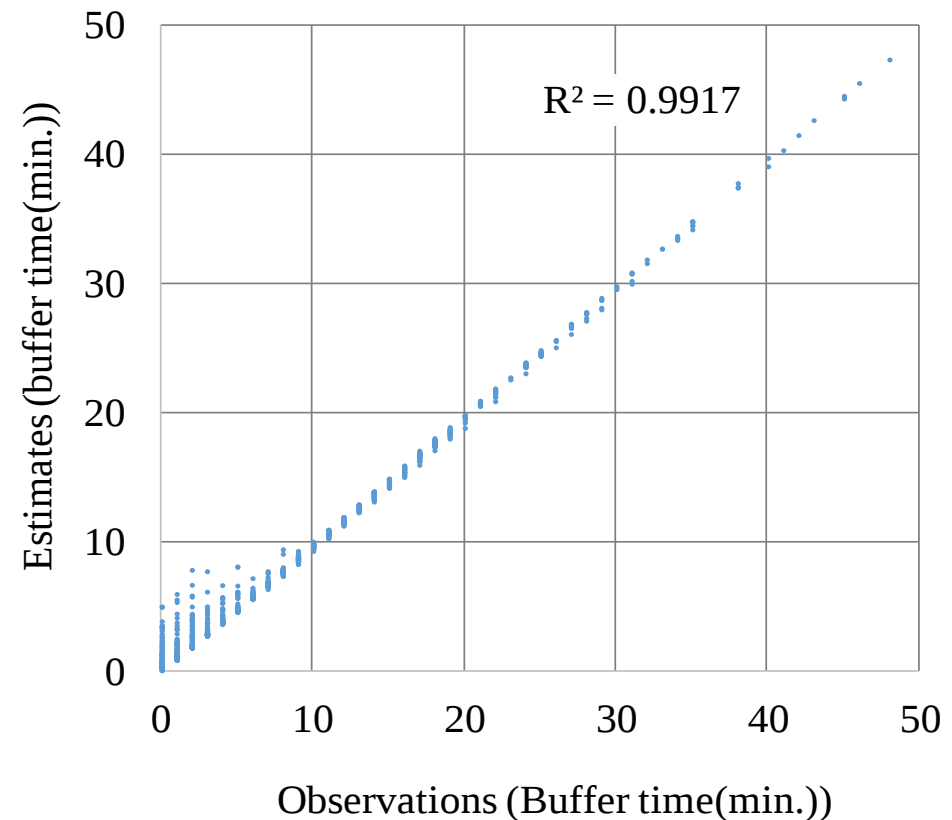
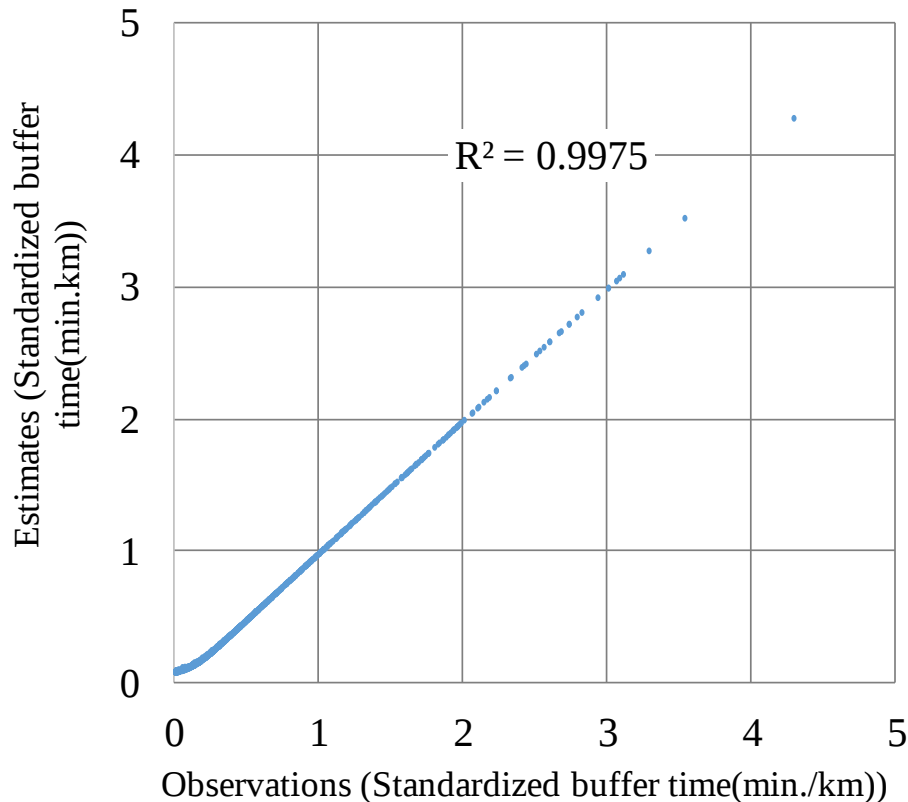
現況再現性の確認

基準化余裕時間

$$\frac{D^* - \mu}{L} = \frac{\ln \frac{\alpha}{\beta}}{\lambda \cdot \exp(\sum_{i=1}^n \theta_i x_i)}$$

余裕時間

$$D^* - \mu$$

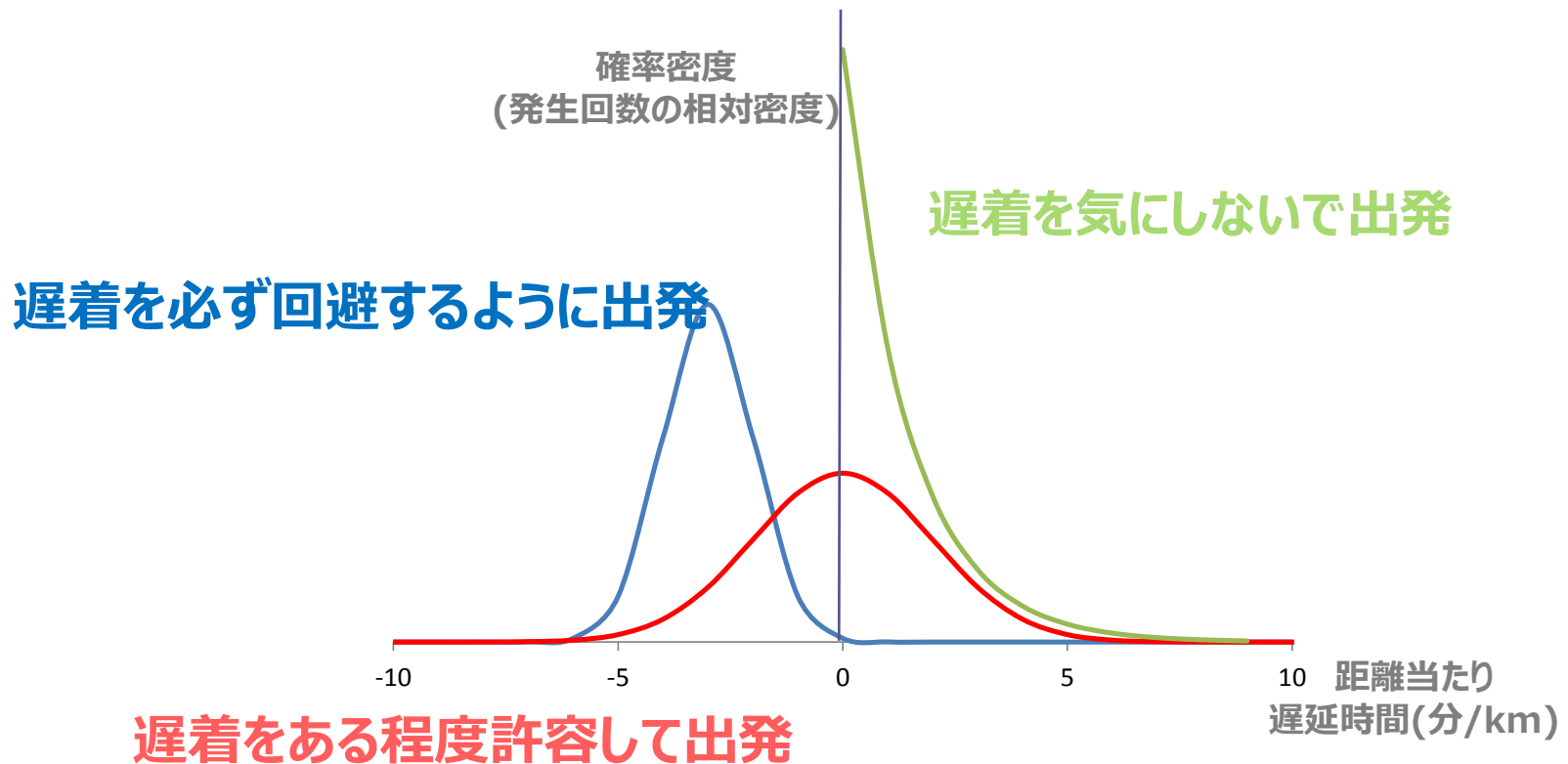


到着時刻分布モデルの推定

-混合分布モデルの適用可能性の検討-

- 高浪裕三:鉄道利用者の到着時刻特性における混合分布適用可能性, 東京電機大学理工学研究科修士論文, 2015.
- 高浪裕三, 高田和幸, 藤生慎:多様性を考慮した到着時刻分布の推定手法に関する研究, 土木計画学研究・講演集, 50, 2014.
- 高浪裕三, 杉山茂樹・藤生慎・高田和幸:鉄道利用者の所要時間分布の特性分析, 土木計画学研究・講演集, 46, 2012.

個々人によって余裕時間設定に差が生じる
→ 到着状況（分布）が多様に



最適到着時刻との乖離の状況に関する調査フォーマット

QA あなたの考える理想の到着時刻から,

(A-1).前後2分以内に到着できるのは, 10回中何回ですか.

(A-2).2分以上遅れて到着するのは, 10回中何回ですか.

(A-3).2分以上早く到着するのは, 10回中何回ですか.

QB QAで, 2分以上遅れて到着する回数は(A-2の回答)とお考えですが, それでは

(B-1).5分以上遅れて到着するのは, 10回中何回ですか.

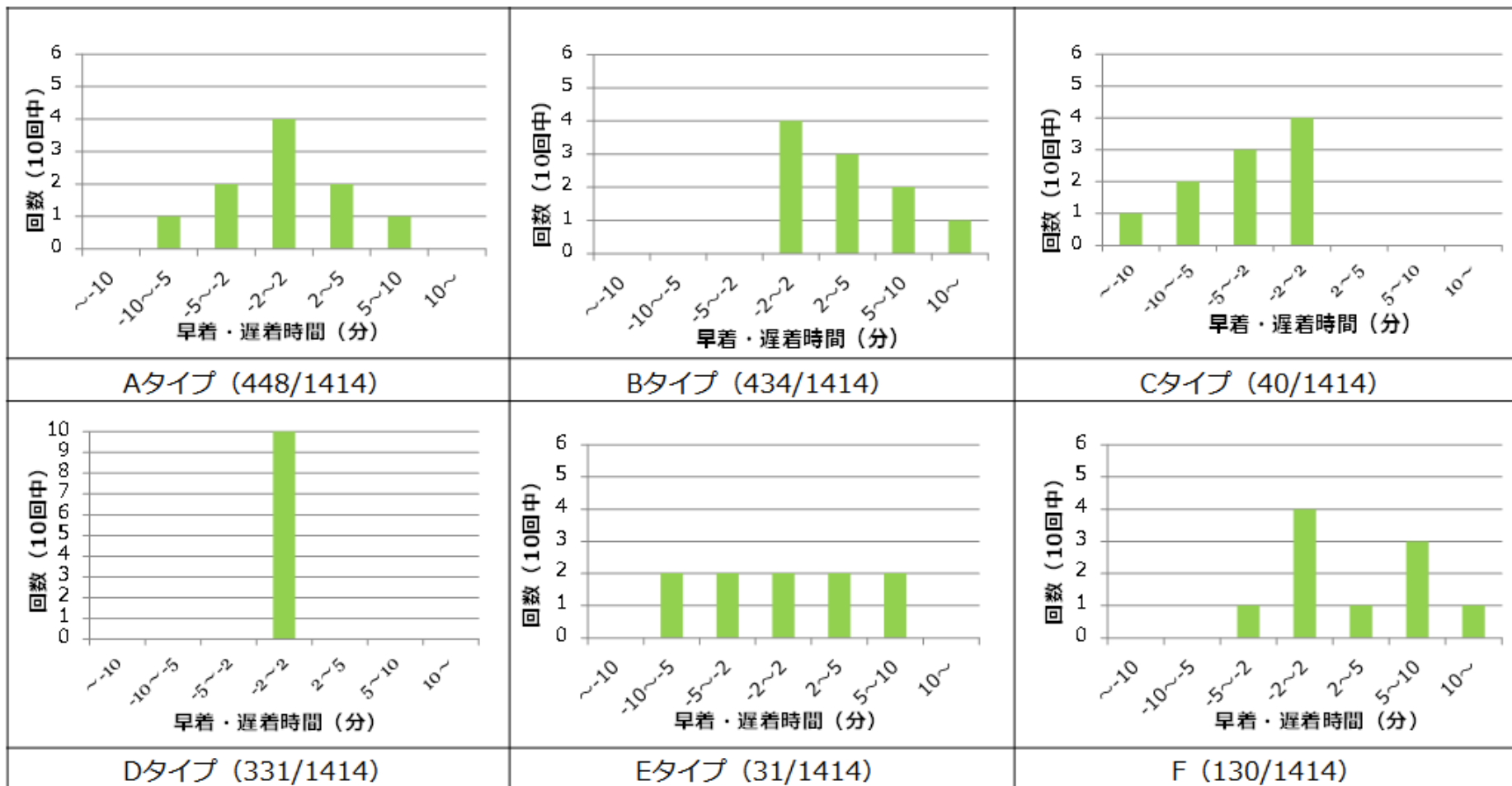
(B-2).10分以上遅れて到着するのは, 10回中何回ですか.

QC QAで, 2分以上早く到着する回数は(A-3の回答)とお考えですが, それでは

(C-1).5分以上早く到着するのは, 10回中何回ですか.

(C-2).10分以上早く到着するのは, 10回中何回ですか.

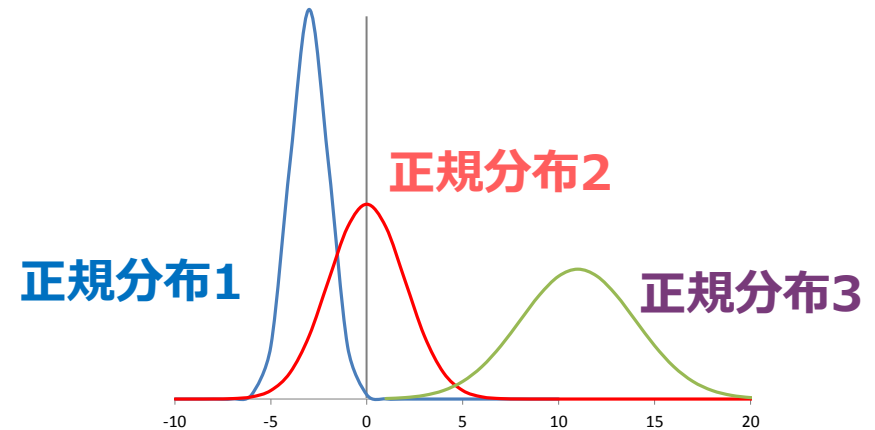
最適到着時刻との乖離状況 →多様なタイプがある



混合モデル

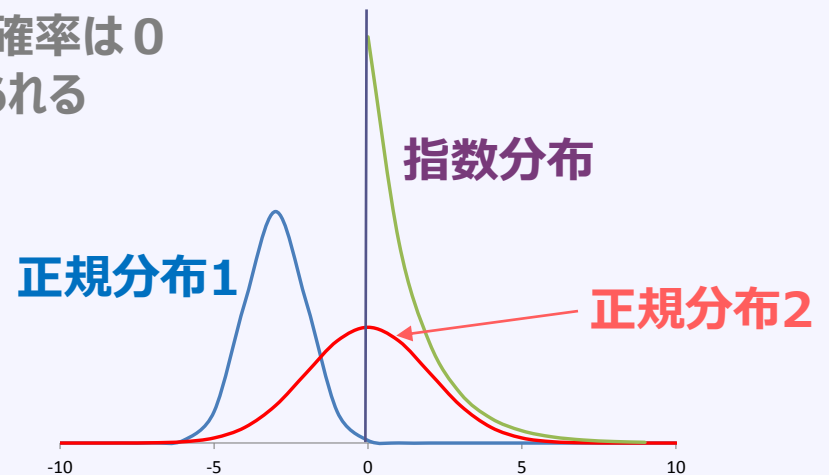
代表的な混合モデル…正規混合モデル

混ざり合った正規分布の
パラメーターを推定するモデル



時刻表を完全に信用して行動した場合，早着する確率は0
→ 到着状況は**指数分布**の形状を取ると考えられる

正規分布と**指数分布**，異なる2種の
分布の混合を説明するモデルが必要



異種混合モデルへ拡張

正規混合モデルの確率密度関数は次のように表される

(パターン認識と機械学習 下 (C.M.Bishop. 著))

混合されている正規分布の数 K

$$p(x|\mu, \sigma) = \sum_{k=1}^K \pi_k N(x|\mu_k, \sigma_k)$$

正規混合分布

k 番目の分布の分担率

平均 μ_k , 分散 σ_k の正規分布

右辺に 指数分布 を加え, 異種混合モデルへ拡張

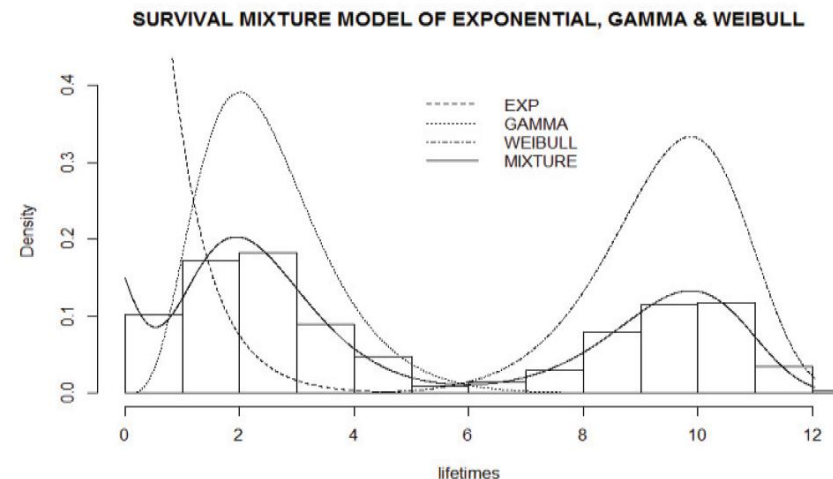
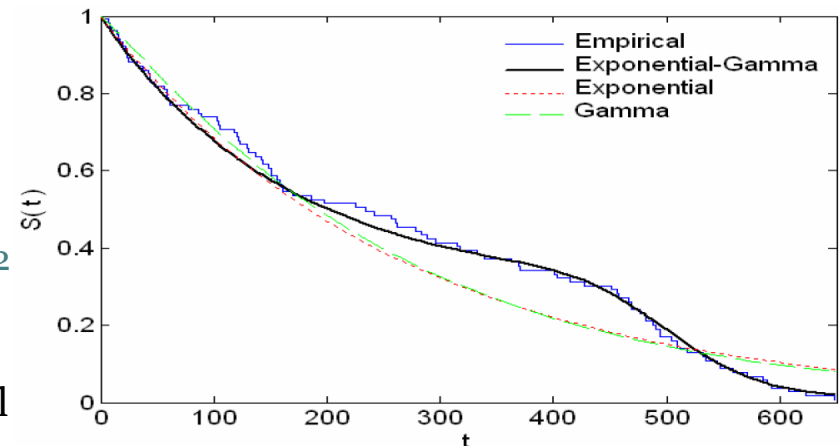
$$\ln p(x|\pi, \mu, \sigma, \lambda) = \sum_{n=1}^N \ln \left\{ \sum_{k=1}^K \pi_k N(x_n|\mu_k, \sigma_k) + \pi_E f(x_n|\lambda) \right\}$$

π_E = 指数分布の分担率
 λ = 指数分布のパラメーター

- EMアルゴリズム(Expectation-maximization)により推定

異種混合分布の研究事例

- Ersioglu and Erol(2010) proposed a mixture model of mixed distributions
 - Modelling heterogeneous survival data using mixture of extended exponential-geometric distributions, *Communications in Statistics - Simulation and Computation*, 39(10), 1939-1952 2010.
- Ülkü Erişoğlu, Murat Erişoğlu and Hamza Erol
 - A Mixture Model of Two Different Distributions Approach to the Analysis of Heterogeneous Survival Data, *International Journal of Computational and Mathematical Sciences*, 5, pp.75-79 :2 2011.
- Yusuf Abbakar Mohammed, Bidin Yatim and Suzilah Ismail:
 - A Simulation Study of a Parametric Mixture Model of Three Different Distributions to Anlayze Heterogeneous Survival Data, *Modern Applied Science*; Vol. 7, No. 7; 2013



異種混合分布の実務への適用（異種混合学習エンジン）

NEC Empowered by Innovation

お問い合わせ サポート・ダウンロード

NECサイト内検索

Japan

Country & Region

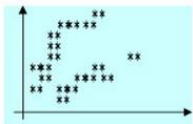
ホーム > ソリューション・サービス > ビッグデータ > ビッグデータを支える先進技術

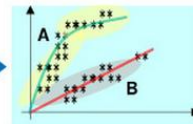
ビッグデータによる価値創造
社会課題を解決する
三位一体のアプローチ

異種混合学習

異種混合学習技術は多種多様なデータに混在するデータ同士の関連性から、特定の規則性を自動で発見するとともに、分析するデータに応じて参照する規則を切り替えます。これにより、“単一の規則性のみを発見して、それを参照するような従来の機械学習”では分析が困難であった「規則性が変化するデータ」でも高精度な予測や異常検出が可能になります。

多種多様なデータから、簡単に複数の規則性を自動で発見し、状況により最適な規則性を選択する、NEC独自のアルゴリズム技術

多種多様なデータ


複数の規則性を自動発見


従来の分析

- 一つの予測式により予測
- 分析専門家が仮説と試行錯誤で、予測式を条件分けして使用

予測式： $Y = a_0X_0 + a_1X_1 + \dots + a_{50}X_{50}$

問題点 専門家が膨大な工数をかけて予測

異種混合学習技術

- 膨大なデータから複数の予測式を自動で導出
- 専門家の仮説なしに、状況（天気など）に応じて参照する予測式を自動で選択

予測式： $Y_n = a_0X_0 + a_1X_1 + \dots + a_{50}X_{50}$

効果 専門がいなくても高精度の予測を維持

想定適用例

エネルギー需要予測ソリューション

ソリューション効果

「異種混合学習技術」で将来の電力需要を正確に予測し、需要に見合った電力の生成および調達を支援

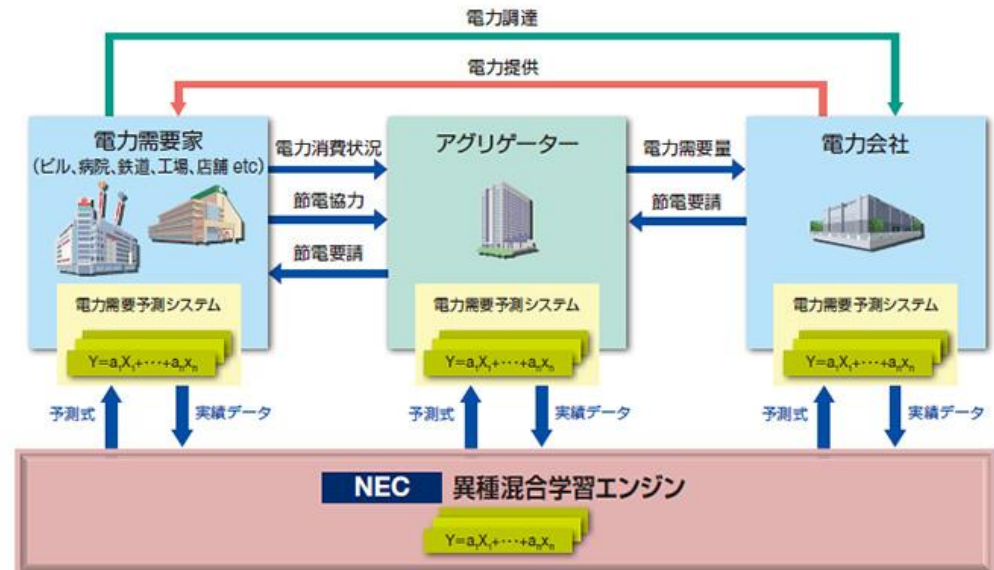
社員の勤務状況や気象情報など、電力使用量に影響する多種多様なデータを、NEC独自の「異種混合学習技術」で分析し、将来の電力需要を高精度に予測することで、省エネや環境に配慮した社会の実現に貢献します。

お客様の声

エネルギーマネジメントを実現し、無理なく無駄を省いて電力利用を実現したい

導入ポイント

- 電力使用量に影響を及ぼす多種多様なデータから電力影響する因子とその重みづけを自動算出
- 手間のかかる複雑なデータ分割作業を手で行うことを実行
- 時間・日・週・月単位で将来の電力需要予測が可能

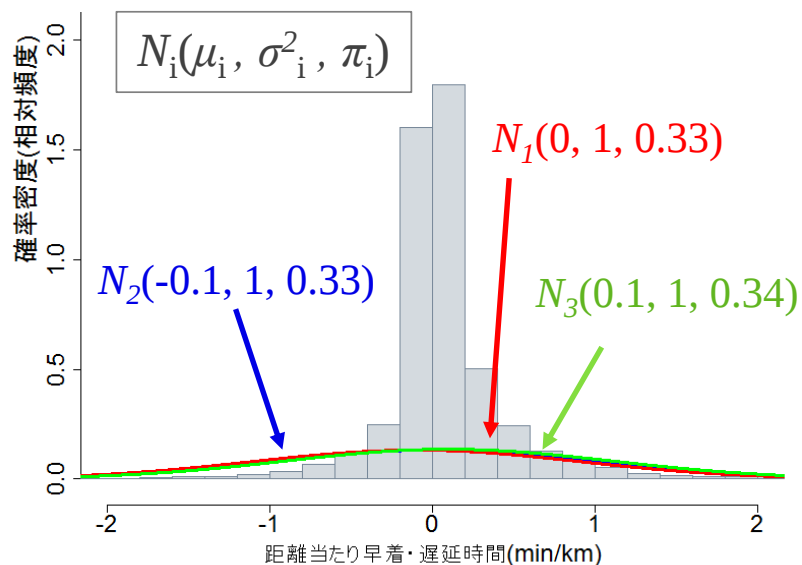


その他の想定適用例

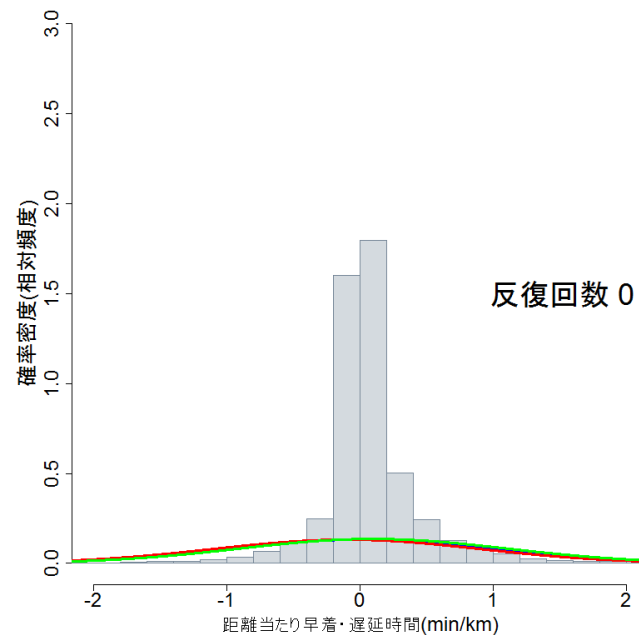
- ヘルスケア：多くの人々の通常時の身体活動データから色々なパターンを自動学習し、パターンから外れる異常状態を検出して、精度の高い健康管理システムを実現。
- 農業：刻々と変化する空気中の成分濃度や土中の物質濃度などを予測し、最適な農作業の内容や時間帯の制御・推薦を行う。

*上記の例に限らず、これまでは整理・分析が極めて困難だった「雑然と収集されたビッグデータ」に対しても、混在するパターン・規則性を分離抽出し、予測や異常検出、制御等に有効活用することが可能。

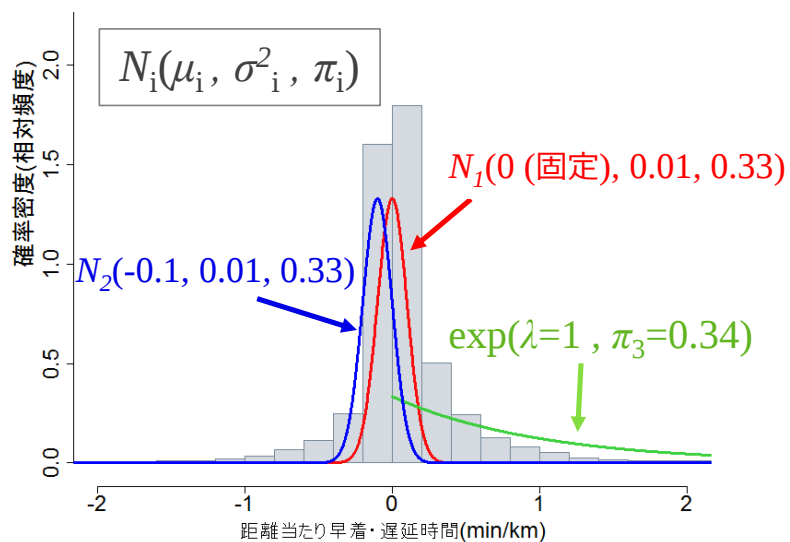
パラメータ推定の条件（正規混合分布）



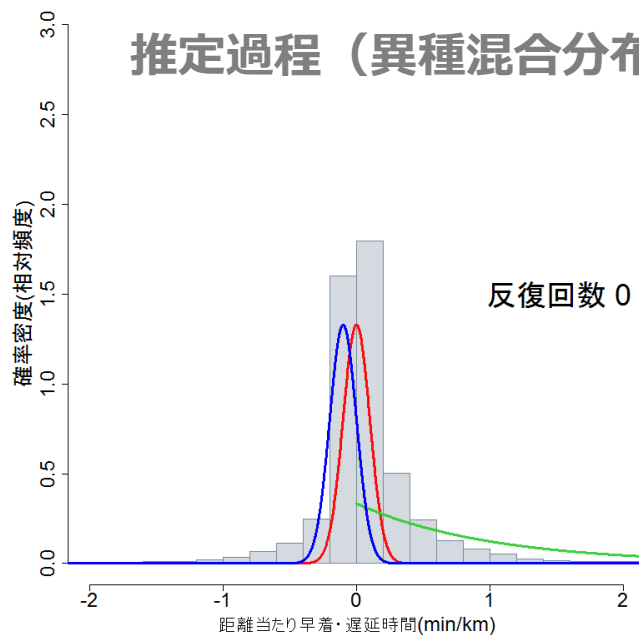
推定過程（正規混合分布）



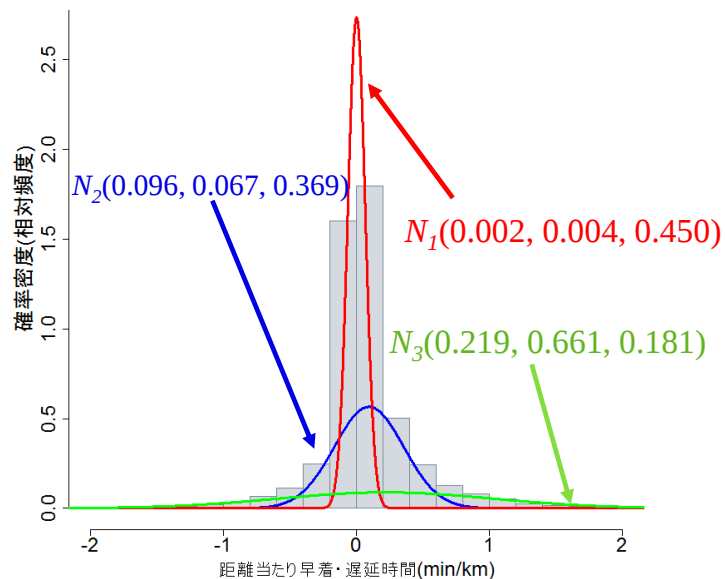
パラメータ推定の条件（異種混合分布）



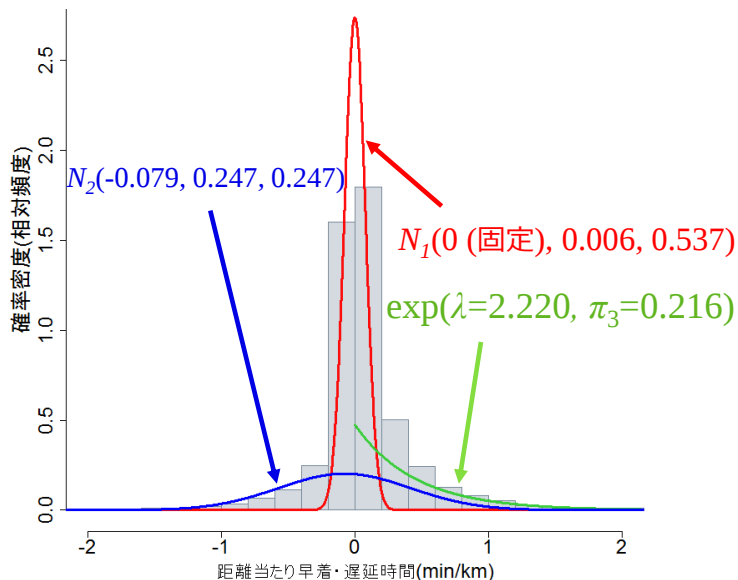
推定過程（異種混合分布）



推定結果（正規混合分布）



推定結果（異種混合分布）

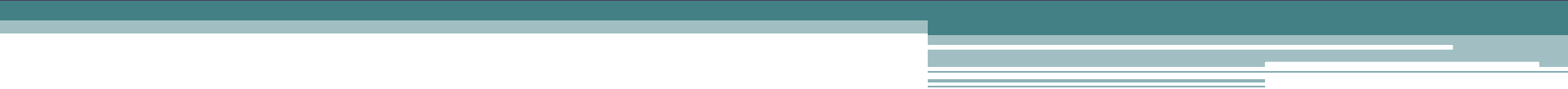


有効性評価 (Bayesian Information Criterion)

分布の要素	対数尤度	パラメータ数	BIC
正規分布1つ	-6594	2	13189
正規分布2つ	-1606	5	3215
正規分布3つ	-1151	8	2307
正規分布4つ	-1166	11	2337
正規分布1つ+指数分布1つ	-4586	4	9174
正規分布2つ+指数分布1つ	-1284	6	2571
正規分布3つ+指数分布1つ	-1120	10	2245

$$\text{BIC} = -2 \times \text{対数尤度} + 2 \times \text{Ln(パラメータ数)}$$

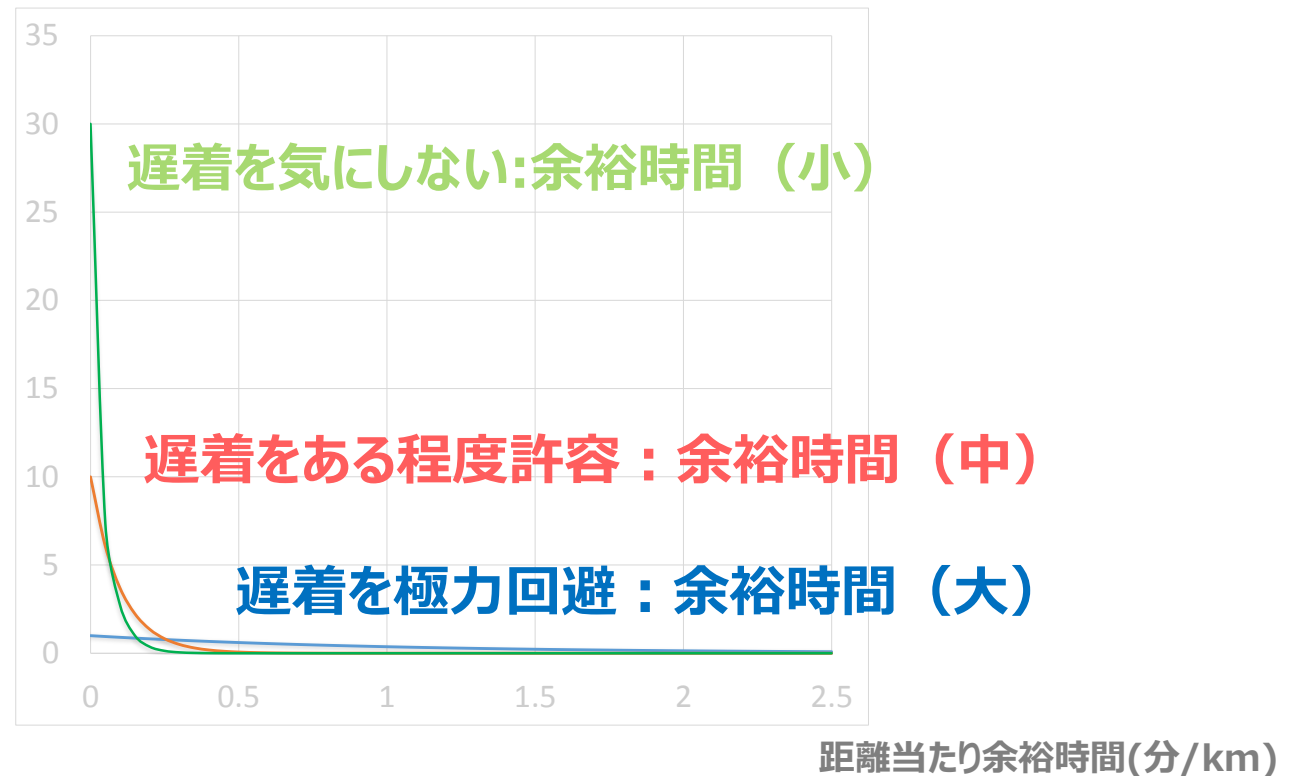
検討課題



出発時刻決定モデルに異種混合モデルを適用（概念）

$$T = \mu + L \cdot T_s$$

$$f(T_s) = \lambda \exp\left(\sum_{i=1}^n \theta_i x_i\right) \cdot \exp\left\{(-\lambda L T_s) \exp\left(\sum_{i=1}^n \theta_i x_i\right)\right\}$$



各分布のパラメータと構成比率が推定できれば，調査サンプル以外への適用も可能となる？