

# 電子メール調査の 悩ましい未来

三根慎二（慶應義塾大学大学院）

2007/1/13

第二回EBAワークショップ

電子メール調査によるエビデンス導出の現状と可能性について

# 発表の概要

1. インターネット調査（特に電子メール調査）  
の概要
2. 図書館情報学分野での事例紹介
3. 物理学研究者を対象とした電子メール調査
4. 電子メール調査から得られるエビデンスとは
5. 提案

# インターネット調査とは

- コンピュータ・オンラインによる調査
- 電子的に調査データ収集を行う調査方式の一つ
- 調査対象者は、主にインターネットユーザ
- 電子調査票による自記式調査
- マルチメディア機能を使う調査票設計が可能

(大隅 2006)

- 郵送調査・自記式の延長線上にある
- 郵送法→電子的手段DBM→電子メール調査→Web調査

(Groves et al 2004)

# インターネット調査の長所と欠点

## ● 長所

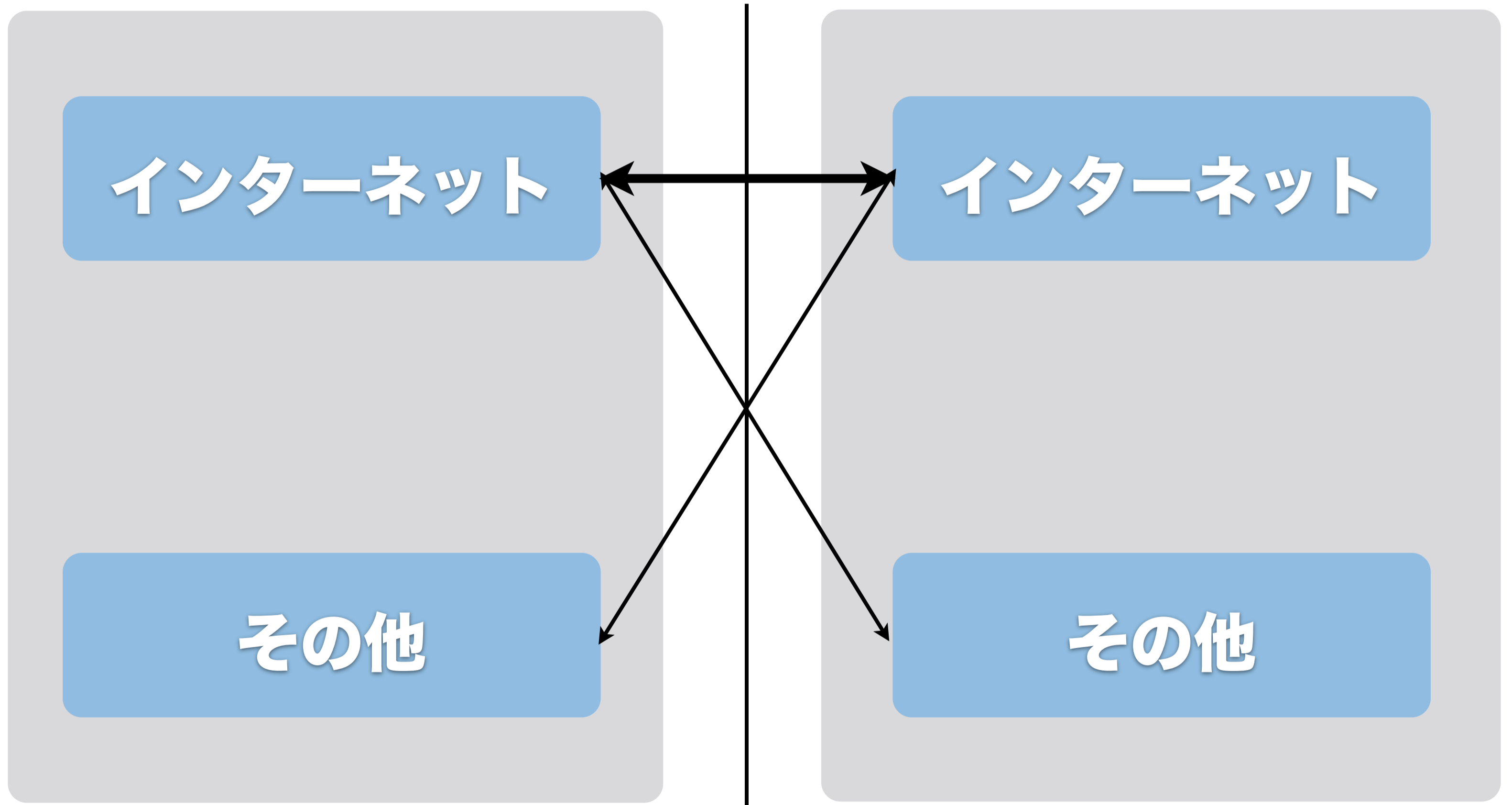
- 簡便性（簡単、使いやすい）
- 速報性・迅速性（早い）
- 調査経費の低減化、廉価性（安い）
- 登録者集団の作り方で回答率が上がる
- マルチメディア対応の調査票設計
- 回答行動の電子的追跡（トラッキング）が可能

## ● 欠点

- 目標母集団が曖昧
- 登録者集団が不透明
- 回答の代表性が不透明（普及率、プロ回答集団）
- 一般に回答率が低い
- 虚偽・代理など不正回答の混入の可能性
- 回答の制御・強制・誘導が起こりうること

# 標本抽出

# データ収集



# インターネット調査の種類

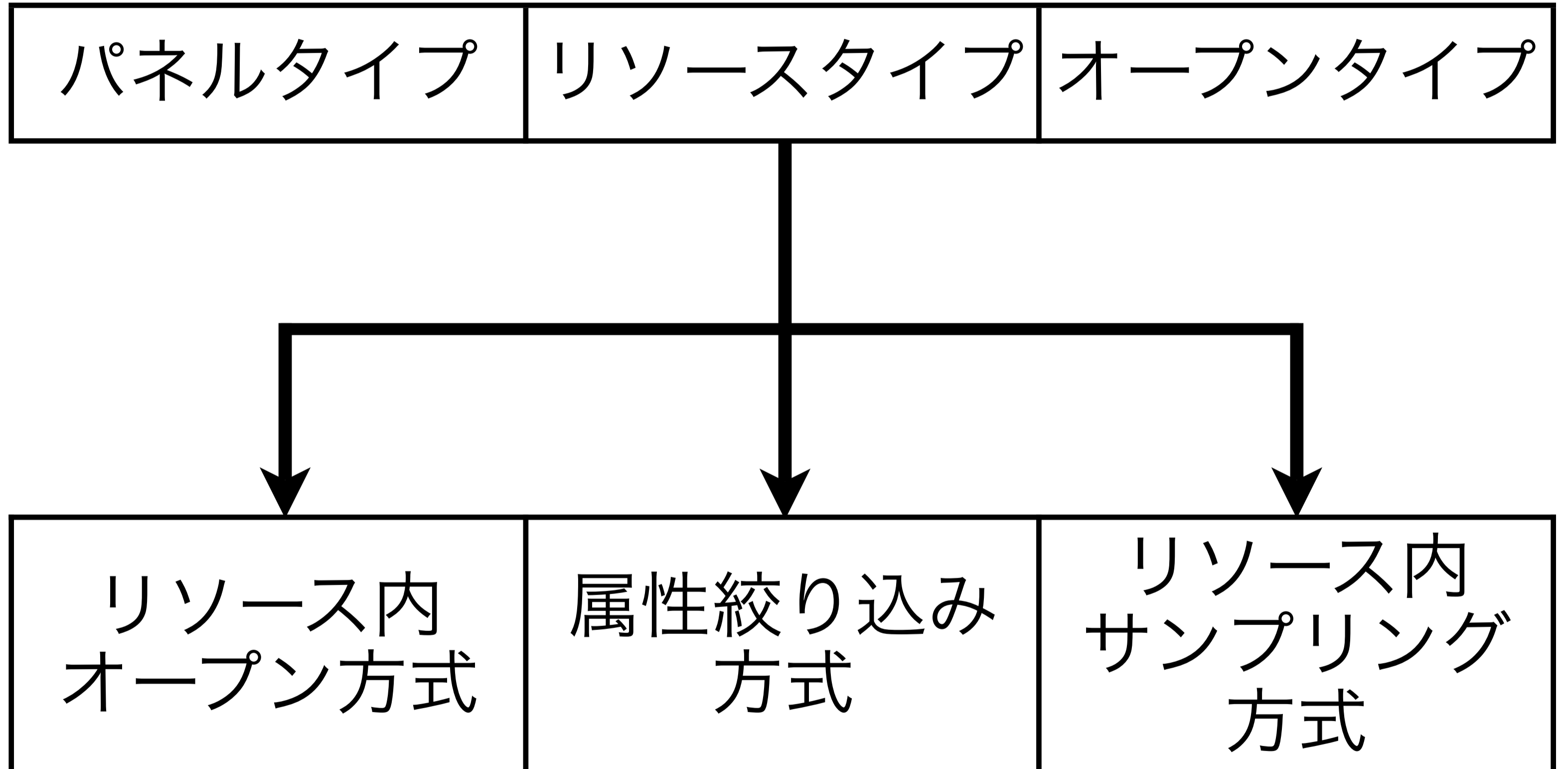
## 確率的アプローチ (科学的)

- インターセプト調査
- リスト型調査
- 混合型調査におけるWeb経由回答
- インターネット利用者の事前収集
- 全母集団からの事前収集

## 非確率的アプローチ (非科学的)

- 娯楽としての投票
- 無制限自己参加型調査
- ボランティア型オプトイン

# インターネット調査の種類



**目標母集団**

**インターネット利用者**

**モニター登録者**

**回答者**

「複数の調査会社にモニター登録して毎週調査に回答している」という回答者が、公募型インターネット調査の主流を占めている (本多 2006)



# 電子メール調査

- **20年の歴史**
  - Kiesler & Sproull (1986)の調査が嚆矢
- **三種類**
  - **量的調査(廣田先生, 三根)**
    - 社会調査
  - **質的調査 (東海林さん)**
    - オンラインインタビュー
    - フォーカスグループ
- **史料**
  - 電子メールのアーカイブ

# 電子メール調査の利点

- 迅速性（調査票の送信と返信）
- 費用対効果（送信費用）
- 繊細な質問・自由回答の分量，率直さ
- 接触困難な人々へのアクセス
- 回答ミスが少ない

*(Schaeffer & Dillman 1998, Roselle & Neufeld 1998, Sheehan 2001)*

# 電子メール調査の問題点

- 母集団の代表性
- 回答率が低い
- 標本抽出枠の獲得が困難
- スпам・ジャンクメールと認識される
- プライバシー・セキュリティ保護が脆弱
- 回答者の技術的なスキル
- データ入力が必要
- 電子メールアドレスの鮮度
- アドレスの存在と実際の利用の不一致

*(Sproull 1986, Ranchhod & Zhou 2001, Sheehan 2001)*

# 電子メール調査の類型(量的調査)

- 電子メールが利用される局面
  - 調査依頼, 調査票の配布, 回答の入力, 返信督促, フォローアップ

調査依頼

電子メール

手紙

調査票の配布

電子メール

WWW

回答の入力, 返信

電子メール

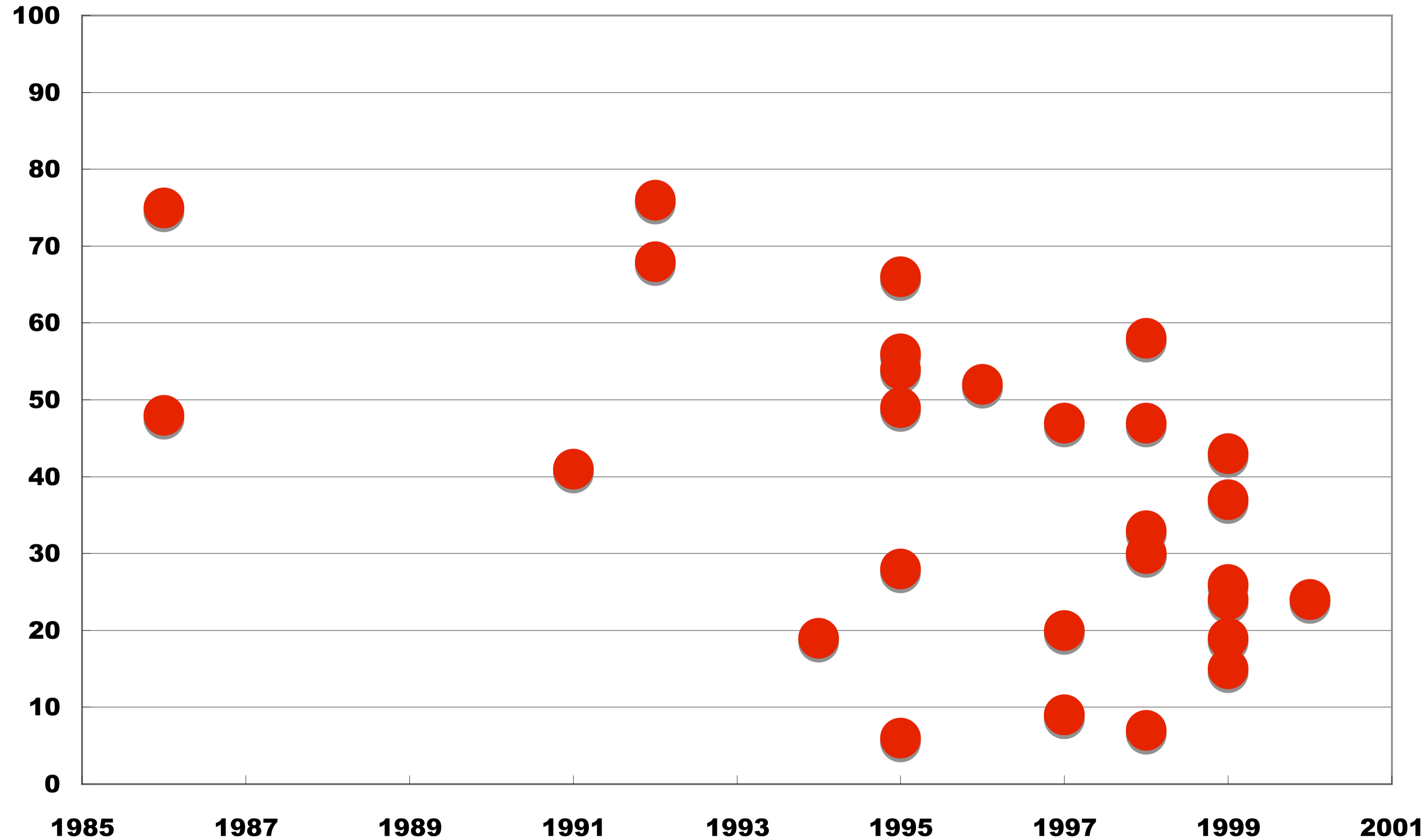
WWW

督促・フォローアップ

電子メール

# 回答率

- **郵送調査と比較して回答率は低い**
  - 平均36.8%（1986-2001）（Sheehan 2001）
  - 同等・高いものもあるが、媒体による違いだけに原因は求められない
- **回答率の計算方法**
  - 非確率的アプローチの場合、回答率を計算できない（信頼性がない）



**回答率の推移** (Sheehan 2001 を元に作成)

# 回答率に影響を与えていた要因

- 調査実施年 (*Sheehan 2001*)
- 調査テーマやサンプルの特徴 (*Sills & Song 2002*)
- 予告状送信
- 催促状送信回数
  - **5.3%～20%上昇** (*Mehta & Sivadas 1995, Smith 1997*)
  - **28.5%(1), 41%(2), 57%(3)** (*Schaefer & Dillman 1998*)
- 個人宛を明記する
- 他形式の調査と併用する

# 他形式との比較

- 同じ調査票を用いて、電子メールから他の手法と同様なデータを入手できるのか？
- 郵送調査などとの比較実験
  - 回答率は低くなる傾向にある
    - 6% vs 27%(Tse et al 1995) 7% vs 52%(1998)
  - 無回答の割合（調査による）
  - 自由回答の文字数が長い

*(Sproull 1986, Bachmann et al 1996, Mehata & Sivadas 1995, Tse et al 1995)*



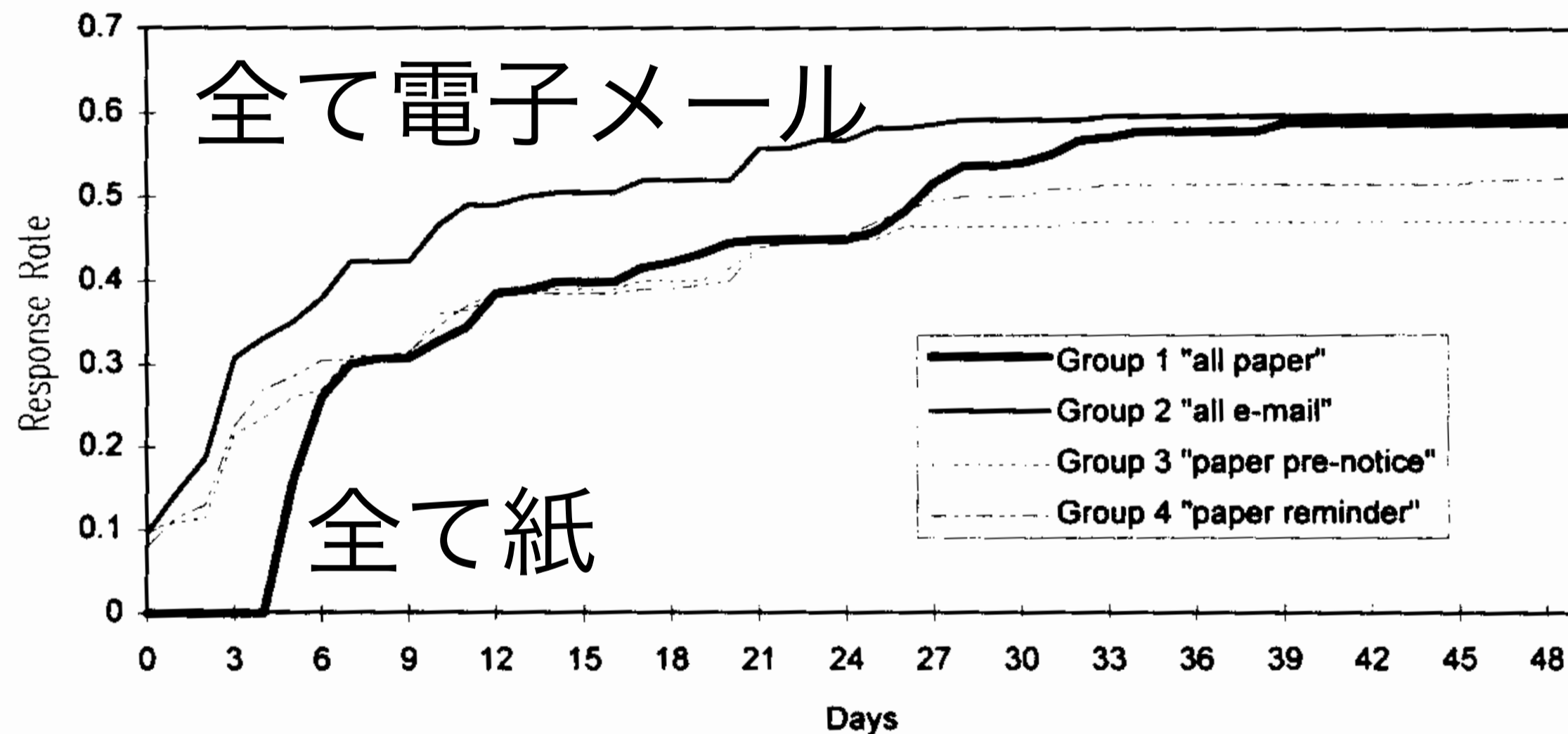
# Schaefer & Dillmanの実験(1998)

	1	2	3	4
事前依頼	紙	E-mail	紙	E-mail
調査票	紙	E-mail	E-mail	E-mail
お礼/督促	紙	E-mail	E-mail	紙
再調査	紙	E-mail	E-mail	E-mail
回答率	57.5	58.0	48.2	55.1

- ワシントン州立大学の教員904名
- 1996年実施
- 「1および2」と「3」とは回答率に有意差あり
- 紙より電子メールの事前依頼をした方が回答率は高い（最初の回答率：45% vs 30%）
- 95%以上の記入率は電子メールの方が高い（69.4% vs 56.6%）
- 44項目中， 30項目が電子メールの方が回答率が高い

# Schaefer & Dillmanの実験(1998)

- 自由回答の記入率も電子メールが高い
- 平均文字数も電子メールの方が長い( 40w vs 10w)
- 平均返信日数：電子メール：9.16 紙：14.39



**Figure 1.** Response rate over time by group (for all groups, the reminder postcard/E-mail was sent on day 7 and the replacement survey was sent on day 21).

出典：Schaefer, D.R & Dillman, D.A. Development of A Standard E-mail Methodology. Public Opinion Quarterly, 62, p.378-397, 1998

# Ranchhod & Zhouの実験(2001)

- 郵送調査とE-mailの比較実験
- イギリスのマーケティング関連の重役
- UK On-line Business Directory, UK Marketing Managers Yearbook 1999から1000名
- 回答率
  - 電子メール：5% 紙：20%
- インターネット関連用語の知識, 電子メール利用の程度に有意差が見られた (電子メール経由の回答者が高い)

# Kaplowitzらの実験(2004)

- 郵送調査とE-mail(www)の比較実験
- ミシガン州立大学の学生19,800名
- 大学から学生の住所とE-mailアドレスを入手

	郵送	郵送 Email	郵送 Email 郵送	E-mail 郵送	Email
回答率	31.5%	29.7%	28.6%	25.4%	20.7%
費用/回答	\$10.97	\$1.31	\$1.80	\$1.53	\$1.32
接触数	4	2	3	2	1

事前に郵送で調査依頼をした方が回答率は高くなっている

# まとめ

- すべての過程を電子メールで行う調査は行われなくなってきた
- 先行研究で指摘されていた利点や長所のうち、その有効性に疑問がふさがるようになってきているものもある
- 時代によって同じ調査課題でも結果が異なるものもある
- E-mailの位置づけが、社会や調査対象者の中で変化していることが伺われる

# 図書館情報学分野の事例

- Eメールを利用した社会調査の事例
- 1990年代前半から登場

*(Savolainen 1998, Bar-Ilan 1999, Zhang 2000)*

- 回答率
  - 0% (Pradhan 1999)~89% (Zhang 2000)
  - (回答数だけのものも多々ある)
- 調査内容
  - 大半が利用者調査
  - 大学図書館の貸し出し方針

# Zhang (2000)の調査

- 図書館情報学分野の学術雑誌8タイトルに掲載された論文の著者203名
- 回答率：89.1%
- 回答経路がWWWか郵送かで
  1. インターネット利用能力の認識
  2. Webの利用頻度
  3. 年齢
  4. インターネット利用年数

に有意差がみられた

# HayslettとWildemuth(2004)の調査

- 北カロライナのレファレンスライブラリアン 300名
- North Carolina Library AssociationとALAの会員リストから無作為抽出
- 回答率：39.0%
- 調査対象者への接触方法（依頼と調査票の配布）のコントロール

	送信数	回答数	有効回答数	回答率
紙 紙	100	53	51	43.2
紙 Web	100	33	28	22.9
Email Web	100	44	39	33.1
合計	300	130	118	39.0



# 物理学研究者を対象とした 電子メール調査

## 概要

- 2003年10-11月実施
- 全世界の研究機関の物理学部所属研究者(教員)
- 電子メールとWWWによる調査票
- 研究者の電子メディア利用の実態, 学術情報  
流通についての意見を聞く
- 回答率: 12.0%

# サンプル構築手順

- **対象母集団**：全世界の物理学研究者
- **枠母集団**：物理学関連の研究機関のWebサイトに掲載されている教員のEmailアドレス
- 全世界の物理学研究機関のリンク集を利用  
(Physnet.netのPhysDep)
- Africa (36), Asia (312), Australia and Oceania (56) Europe (726), North America (522), South America (97) (2007.1現在)
- 合計で、20,619名のE-mailアドレスを収集
- 2003年8月から9月
- **計画標本**：プレテストの100名を除いた残りから無作為抽出で3,000名

# 調査票

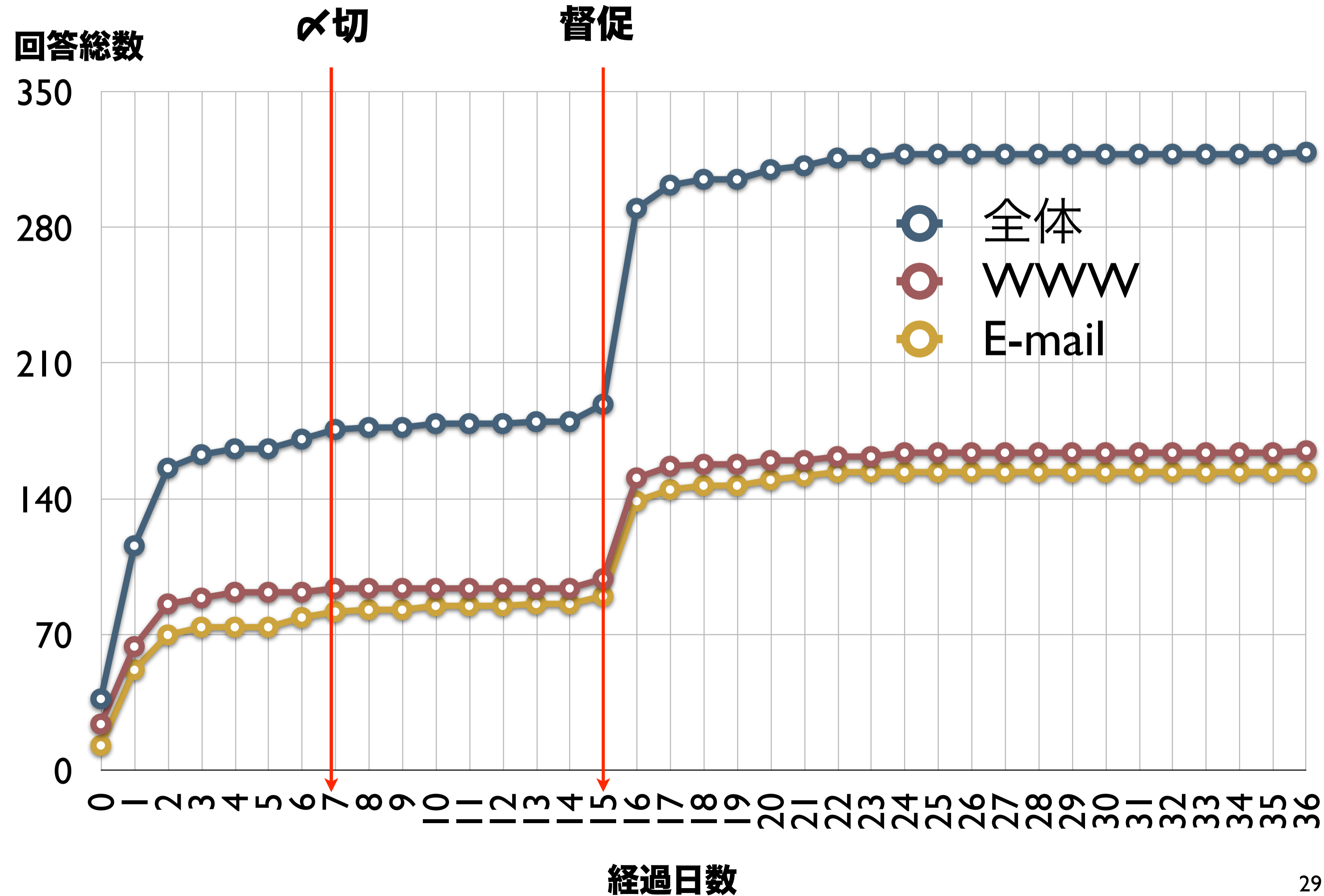
- E-mailの本文とWebページに掲載
  - 内容は同一（Webには説明はなし）
- 合計21問
  1. 調査依頼(E-mailのみ)
  2. フェイスシート
  3. 論文の執筆
  4. 論文の公開
  5. 電子メディアの利用
  6. 学術情報流通についての意識・態度

# 回答率

送信数		2978	-
有効送信数		2666	100%
有効回答数	WWW	165	6.1%
	E-mail	154	5.9%
	計	319	12.0%
各項目の無回答率		0.3%~11.6%	

\*IPアドレスからみると、重複回答はなかった

# 経過日数と累積件数



# 回答者の属性

偏りはあるか

年齢		
60代	58	18.5%
50代	77	24.6%
40代	88	28.1%
30台	77	24.6%
20代	13	4.2%
計	313	100.0%

職位		
教授	141	44.3%
準教授	70	22.0%
助教授	35	11.0%
講師	34	10.7%
その他	38	11.9%
計	318	100.0%

# 回答者の属性

地域	標本		回答者	
	数	割合	数	割合
欧州	1618	54.3%	181	56.7%
北米	602	20.2%	71	22.3%
アジア	311	10.4%	26	8.2%
南米	291	9.8%	23	7.2%
豪州	123	4.1%	17	5.3%
アフリカ	33	1.1%	1	0.3%
合計	2978	100.0%	319	100.0%

e-mail:TLD www:IPアドレス で判断

# 回答者の属性 電子ジャーナルの利用頻度(参考)

	実施年	分野	EJ利用率	形式
三根	2003	物理	89.5%	E-mail
倉田ら	2003	物理	84.7%	混合
倉田ら	2003	化学	90.0%	混合
倉田ら	2003	病理学	87.1%	混合
Bar-Ilan	2002	自然科学	99.5%	E-mail



# 回答者の属性 arXivの利用頻度(参考)

	実施年	分野	arXiv利用率	形式
三根	2003	物理	71.5%	E-mail
松林ら	2003	物理	31.7%	混合
ALPSP	2000-2001	物理	32%/21.6%	E-mail
倉田ら	1999	物理	33.1%	紙
Lawal	-	物理天文	51.6%	E-mail
高島	1998	物理	81.3%	E-mail

# 初期回答者と後期回答者

- 初期回答者：督促なしで回答 N=180
- 後期回答者：督促後回答 N=139
- 初期回答者の方が、全体平均および後期回答者よりも電子メディアを頻繁に利用している傾向
- 後期回答者は、全体平均よりも低い傾向
- ただし、有意差があるわけではない

# 自由回答

- 全回答中28件 (8.8%) (n=319)

	WWW	E-mail
件数	17	11
文字数	4,791	317
率直さ	◎	△

# 過激な反応

- 調査に私の時間やお金を費やしたくない
- 全くナンセンス
- 今度送ったらスパム扱いにする
- 人の手を煩わせる前に、もっと回答しやすいインターフェースを用意したら？
- この無礼をやめて、私のアドレスを忘れ  
くださいませんか？
- Spammer and MINE Shinjiで検索

# まとめ

- **回答回収時間は確かに早く， 経済的**
  - 最短は送信から12分後， 4日で50%
  - 費用は実質0
- **回答率の低さ**
  - 未達が標本の1割以上
  - 督促は効果があるが， 5%程度
  - Eメール調査だからといって， 年齢に若年層が多い， 利用頻度が高いといった偏りは見受けられない（arXivは除く）
- **無回答誤差はわからない**
  - テーマに関心がある人が回答している可能性

# 電子メール調査から得られる エビデンスとは

## 電子メール調査はかなり制限的な調査手法である

- 世論調査など，一般母集団の調査には向いていない
- 電子メール版住民基本台帳，選挙人名簿はない

## 調査環境の悪化

- 電子メールで見知らぬ他者への接触に大きく期待できる時代ではない

## 電子メールを使うことの意味

- 費用や迅速性から選択することは乱暴
- MLなどでの参加呼びかけ・自己参加は信頼性がない

# インターネット普及率

- **情報通信白書 (H18)**

- 8,529万人(2005年)
- 人口普及率：66.8% 世帯普及率：87.0%
- 12歳以下, 50, 60年代の利用率は全体と比較して差が大きい
- 年収600万円未満は影響度がマイナス

- **インターネット白書(2006)**

- 7,361万9千人(2006.2)  
人口普及率：57.6% (推計) 世帯普及率：57.3%  
世帯浸透率：85.4%
- 広く普及しているのは事実であるが、国民全体を反映しておらず母集団（インターネット利用者）を把握する確かな方法もないのが現状

# 調査誤差

- カヴァレッジ誤差
  - Eメールのリストの鮮度  
(8~28%の未達 Sills & Song 2002)
- 標本誤差
- 無回答誤差
  - 一般的に低い回答率
- 測定誤差
  - 電子的な調査票のデザイン
- 集計誤差
  - データの検証方法



# 電子メール調査が可能な環境

- 対象母集団が、確実にインターネット利用者でなおかつ日常的に利用している
- 混合型調査の一部に利用
- 事後調査やインタビュー調査など質的なデータの入手に利用
- 調査票がマルチメディア要素を含む
- 調査票が繊細な質問を含む，自由回答が多い
- 電子メールでしか接触できない集団
- 「魚群探知型調査」(平松 2006)で十分な場合
- (費用・時間がない場合)

# *The Checklist for Reporting Results of Internet E-Surveys(CHERRIES)*

- 調査設計
- IRBの承認・インフォームドコンセント
- 開発とプリテスト
- 採用過程, 調査票にアクセスするサンプルの記述
- 調査の管理
- 回答率
- 多重回答の回避方法
- 分析手法

# 電子メール調査に必要なもの

## 良い調査票の作成が前提

### 回答率を上げるために

- 事前検証（プリテスト， 技術的確認）
- 電子メールあるいは郵送での簡潔な依頼状
  - ただし， オプトイン
- 複数回の接触・督促

### 回答の代表性を高めるために

- 通常の標本抽出枠を用い， その中から無作為抽出を行う
- 目標母集団が電子メールを使っており， 電子メールアドレスのリストがある
  - 大学教員， 政府職員， 企業職員， 専門組織（Dillman 2000）
- 幅広い年代， 職業を対象としない

# 提案

- **エビデンスデータベース**
  - 調査方法, 調査対象, 研究志向・実務志向などで検索可能にする
  - 調査票の共有 (電子ジャーナルの時代)
- **インターネット調査の記述の指針**
  - 論文に記述すべき項目の標準化
    - 母集団, 標本, 抽出手順, 回答率
    - e.g. Journal of Medical Internet Research
  - メタ分析を可能にするために

# 調査票デザインの原則 (大隅ほか 2003)

1. Welcome Screenを設ける
2. 第1問は1スクリーン、すべての回答者が容易に理解できるもの
3. 従来の紙による質問紙に似たフォーマット
4. 1行は短く
5. 必要な操作の説明
6. 操作の説明は質問ごとに(冒頭一括は×)
7. 回答の強制はしない(答えなければ次に進めないのは×)
8. 分岐が必要なとき以外はスクロールタイプを
9. 1つの設問は1つのスクリーンに収まるように
10. あとどれくらいで調査が終了するかが分かるように
11. 「あてはまるものすべて」や自由記述には気をつけること

1. 調査の実施時期
2. 枠母集団
3. 計画標本の抽出方法
4. 計画標本数及び回収標本数
5. 調査票
6. 調査方法
7. 単純な平均や合計以外の集計を行った場合はその方法
8. 枠母集団の作成方法及び管理方法
9. 計画標本の抽出方法
10. 回収打ち切りのタイミング
11. データ入力のチェック方法
12. データ・クリーニングの方法及び自動修正の件数