




集団における習慣的な食品摂取量の 分布を推定するために考慮すべき事項

国立保健医療科学院

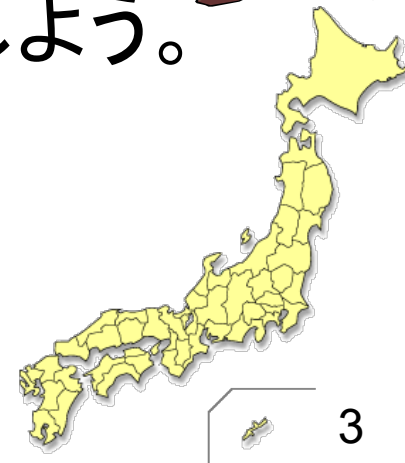
生涯健康研究部

横山徹爾

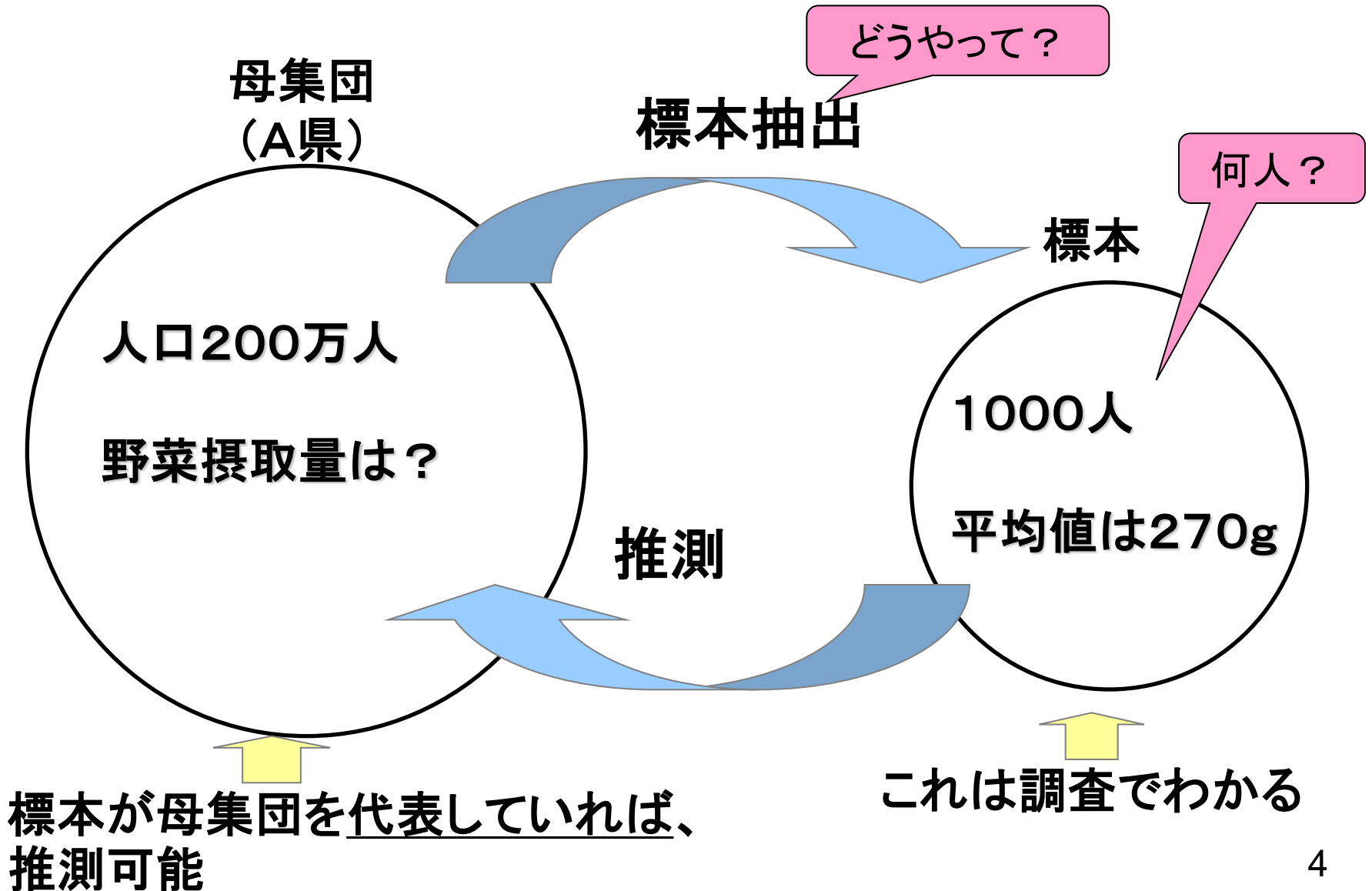
- 
- ✓ 集団代表性のある標本抽出について
 - ✓ 標本調査の誤差について
 - ✓ 測定誤差について
 - ✓ 人口構成による重み付け
 - ✓ 習慣的な摂取量の分布推定
 - ✓ 体重あたり摂取量の分布について

標本調査の目的

- 国民全体（母集団）における食品摂取量の分布を知りたい。
- 全員調べれば（悉皆調査）わかるが、とうてい不可能である。
- そこで、一部の人たち（標本）を調べて、国民全体（母集団）の様子を推測しよう。



母集団と標本



無作為抽出

- 調査対象としている人口全体が**母集団**。例えば、国民健康・栄養調査では日本人全体が母集団である。
- 標本抽出を行う場合は、母集団をいくつかの**抽出単位**（“個人”，“世帯”，“**国勢調査区**”，“**単位区**”など）に分け、全ての抽出単位が選ばれる確率が等しくなるように工夫する（簡単に言うとランダムに選ぶ）。これを**無作為抽出**という。
- 母集団の特性を推定するためには、無作為抽出を用いなければならない。

クラスター抽出法

- A県内の国勢調査区から、乱数によって選んだ10地区の住民全員を対象として調査を行う、というように、母集団をいくつかの集落＝**クラスター**（**国勢調査区**、**単位区**など）に分け、**クラスターを抽出単位として無作為抽出**を行い、選ばれたクラスター内の**構成員全員**を調査対象とする方法。
- 調査地域が広く（例えば全国、全県レベル）、**訪問が必要**な調査で行われることが多い。
 - **長所**：調査のための移動・訪問の手間を減らせる。
 - **短所**：単純無作為抽出に比べて、少し精度が劣る。
- 国民健康・栄養調査は、都道府県別のクラスター抽出（**層化クラスター抽出**）によって実施されている。 6

✓ 集団代表性のある標本抽出について

 ✓ 標本調査の誤差について

✓ 測定誤差について

✓ 人口構成による重み付け

✓ 習慣的な摂取量の分布推定

✓ 体重あたり摂取量の分布について

標本調査と誤差・偏り

- 誤差

- 真の値と、観察した値とのずれ。標本調査ではつきもの。

- ランダム誤差

- 偶然現象によって生じたずれ。標本抽出による誤差を特に標本誤差という。統計学で**誤差の大きさがわかる**。

- 系統的誤差

- 何らかの理由により、一定方向（正または負）に生じたずれ。「**偏り**」ともいう。統計学でも（基本的に）**どの程度のずれがあるのか分からない**。

標本誤差と測定誤差

- **種類**によって、

 - 標本誤差**・・・標本抽出の段階で生じる誤差

 - 測定誤差**・・・測定の段階で生じる誤差

- **性質**によって

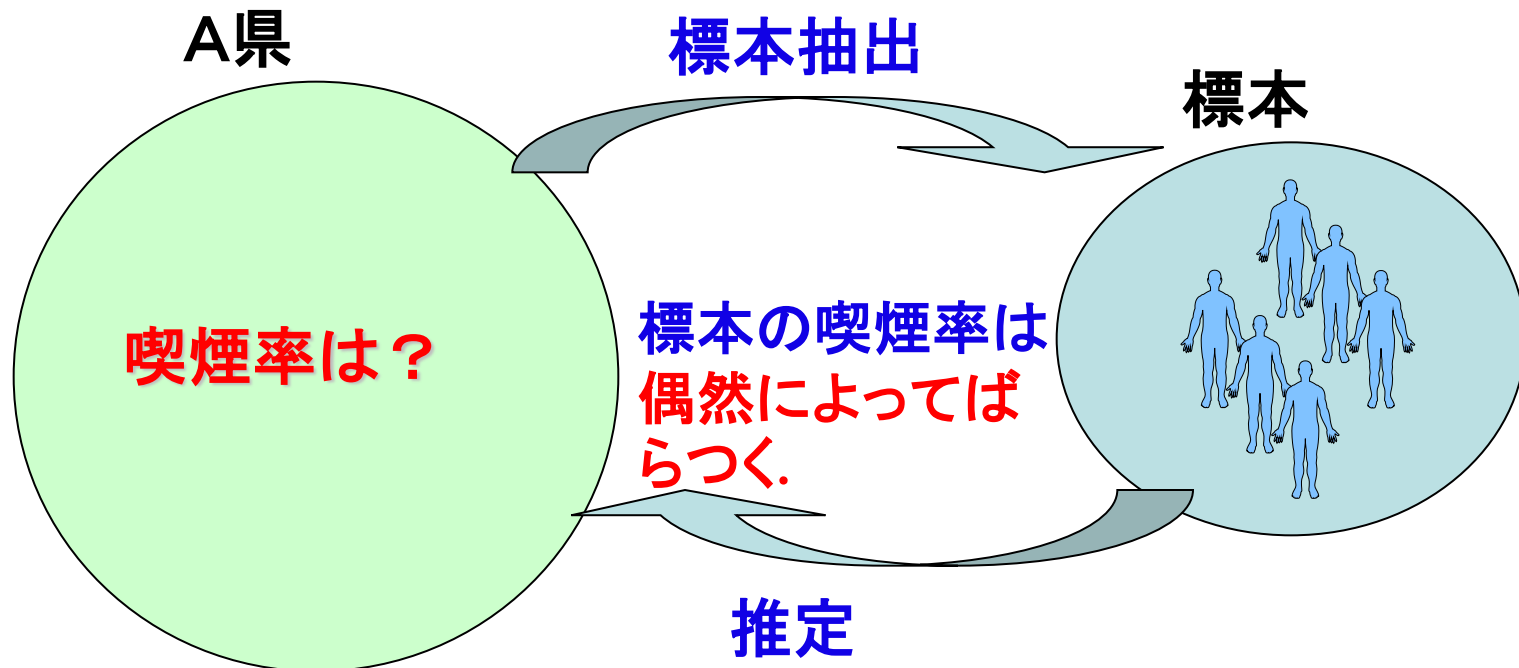
 - 系統的誤差**・・・測定値が真の値から特定の方向にずれている(偏り)

 - ランダム誤差**・・・“ずれ”が、特定の方向に偏らない(平均するとゼロ)

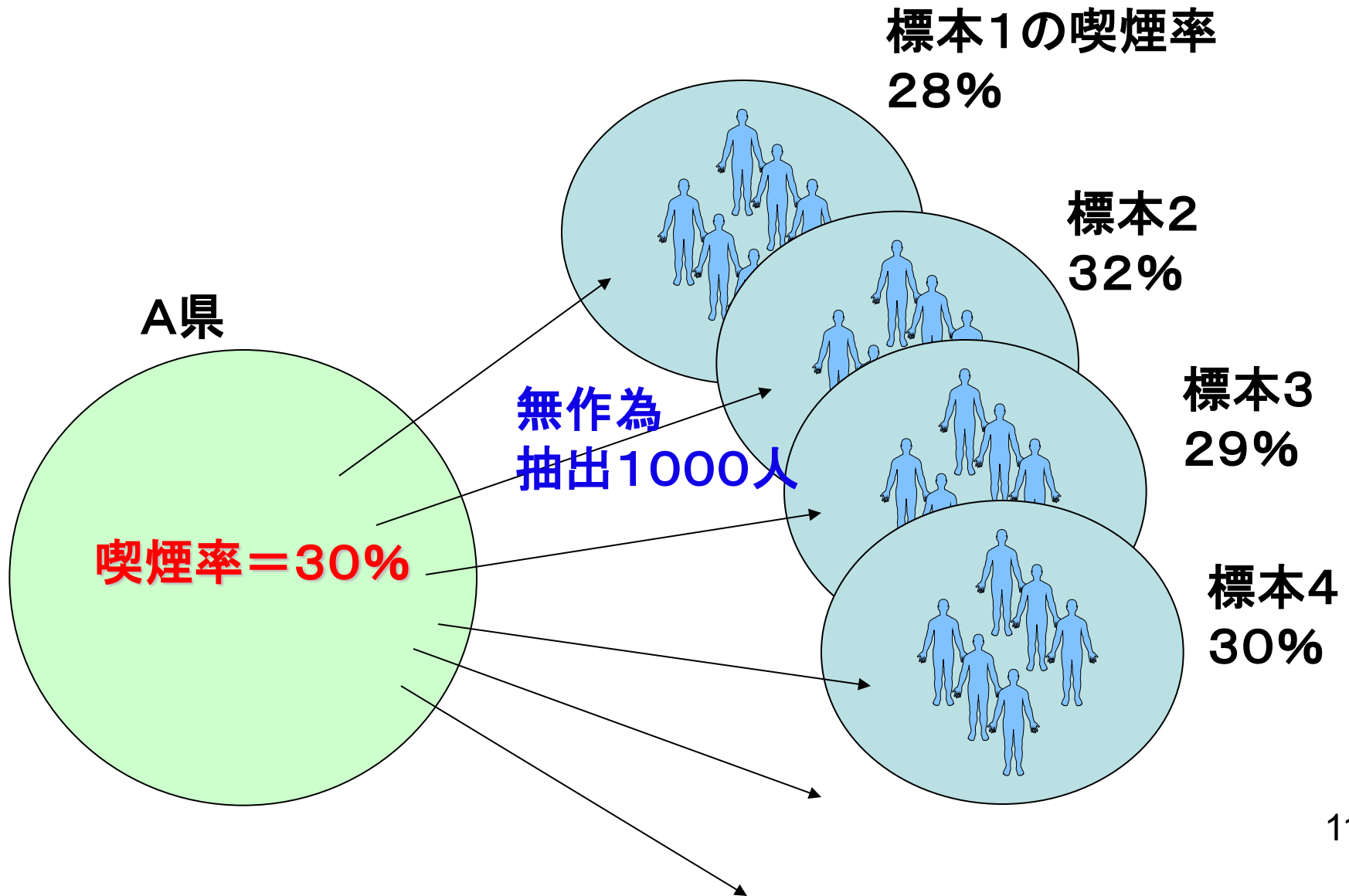
ランダム誤差(標本誤差)

●ランダム誤差

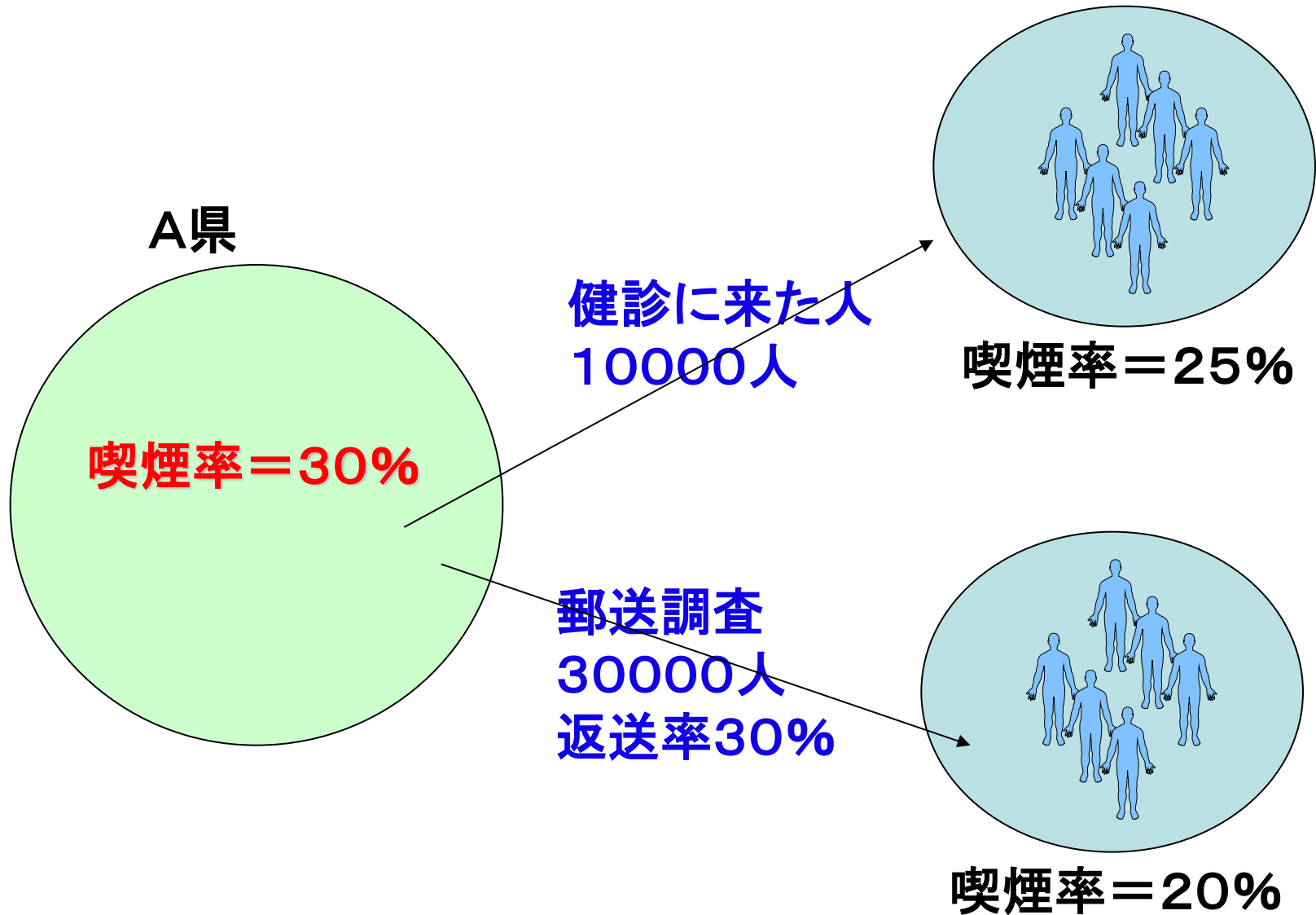
偶然によっておこる。ランダム誤差の大きさは、**標本数が多いほど小さい**。正負方向に同様に起こり、平均するとゼロになる。



ランダム誤差(標本誤差)の概念



系統的誤差(標本の偏り)の概念



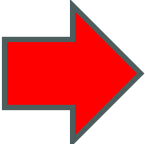
- ランダム誤差を減らすには

- 調査人数を増やす。

- 系統的誤差(偏り)を減らすには

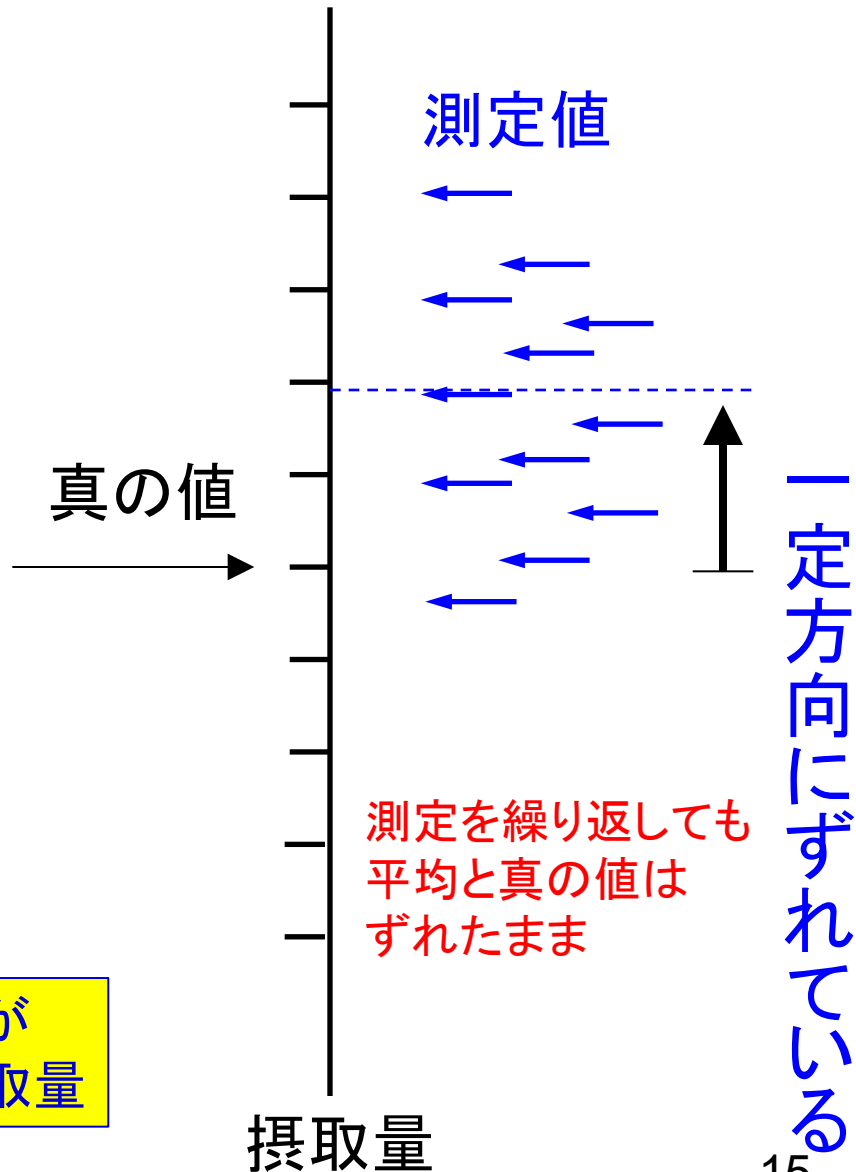
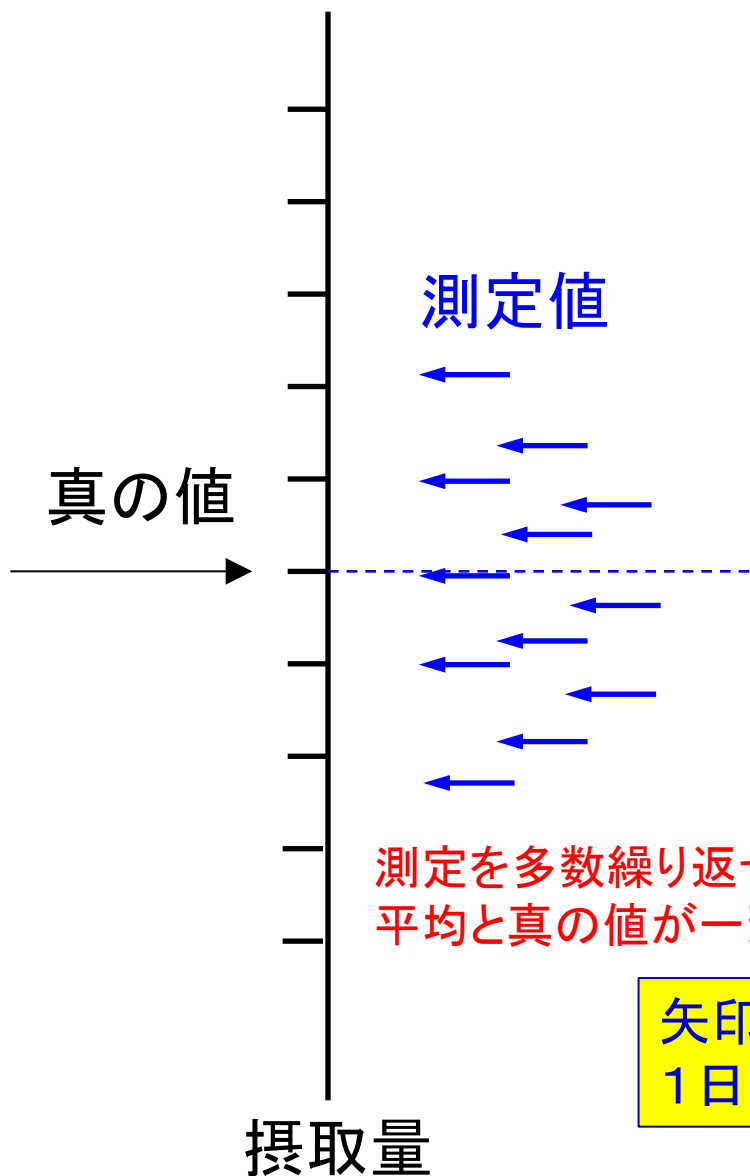
- 無作為抽出法を用いる。

- 回収率を高める。(回収率が低いまま人数を増やしても偏りは減らない)

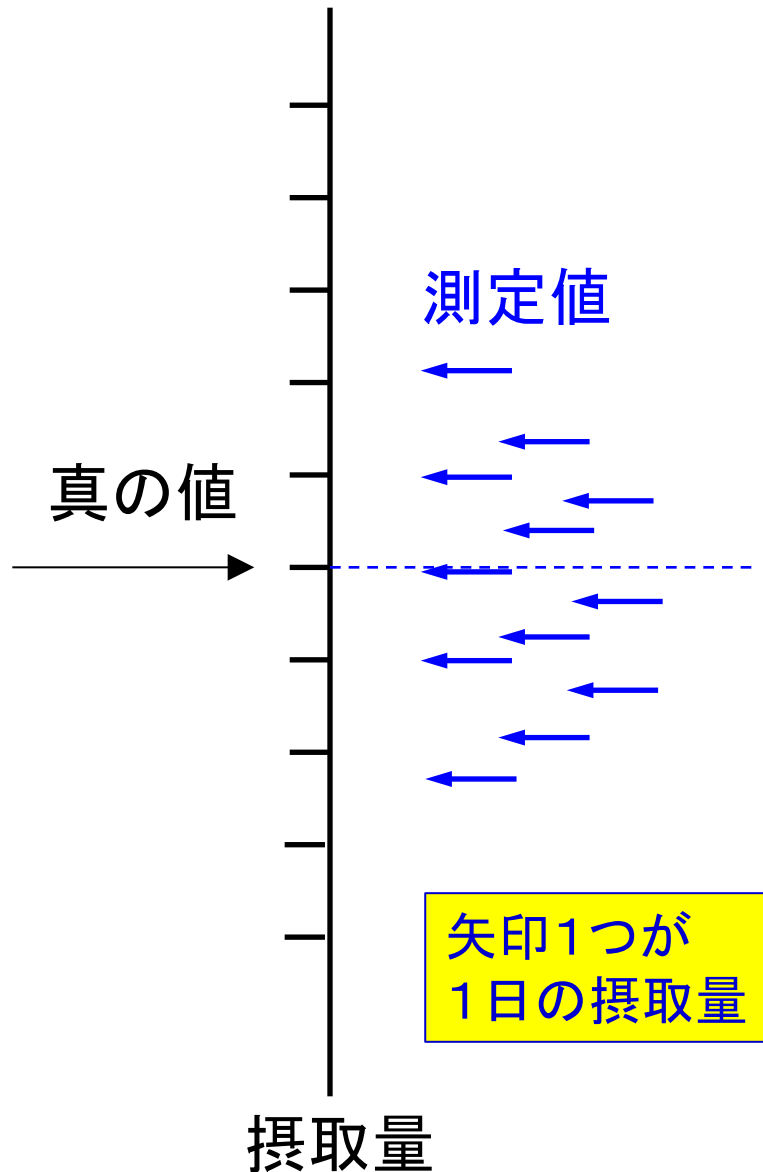
- ✓ 集団代表性のある標本抽出について
- ✓ 標本調査の誤差について
-  ✓ 測定誤差について
- ✓ 人口構成による重み付け
- ✓ 習慣的な摂取量の分布推定
- ✓ 体重あたり摂取量の分布について

ランダム誤差

系統的誤差 (とランダム誤差)



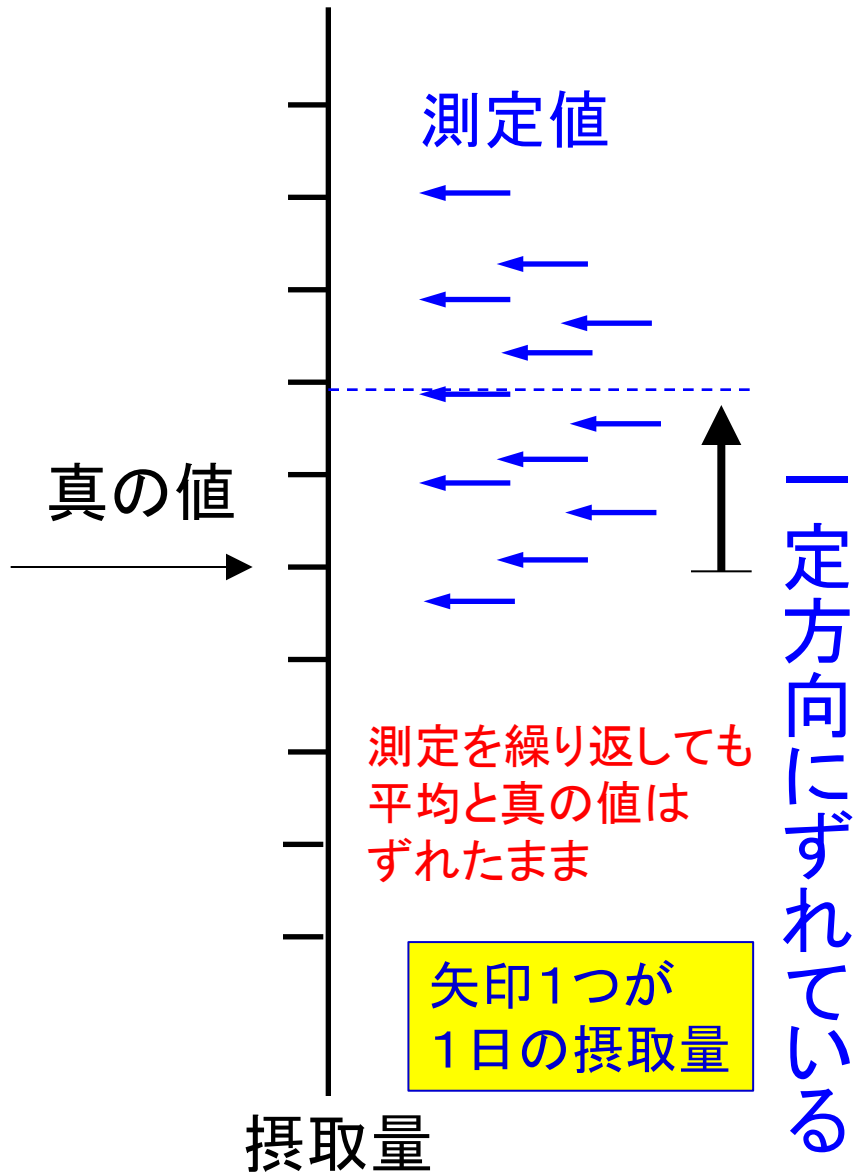
ランダム誤差



測定のランダム誤差の例

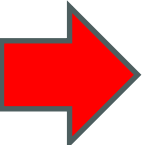
- 日間変動、季節間変動
 - 食事(食品)摂取量は、毎日異なる。
 - ✓ 複数日測定すれば、その平均が真の値(長期間の平均的な摂取量)に近づく。
 - ✓ 例えば、同じ人に3日間食事調査を行ってその平均値をとる。
 - 季節によっても異なる。
 - ✓ 季節毎に複数日調査を行う。
 - ✓ または、調査の欠点を理解したうえで解釈する。

系統的誤差 (とランダム誤差)



測定の系統的誤差の例

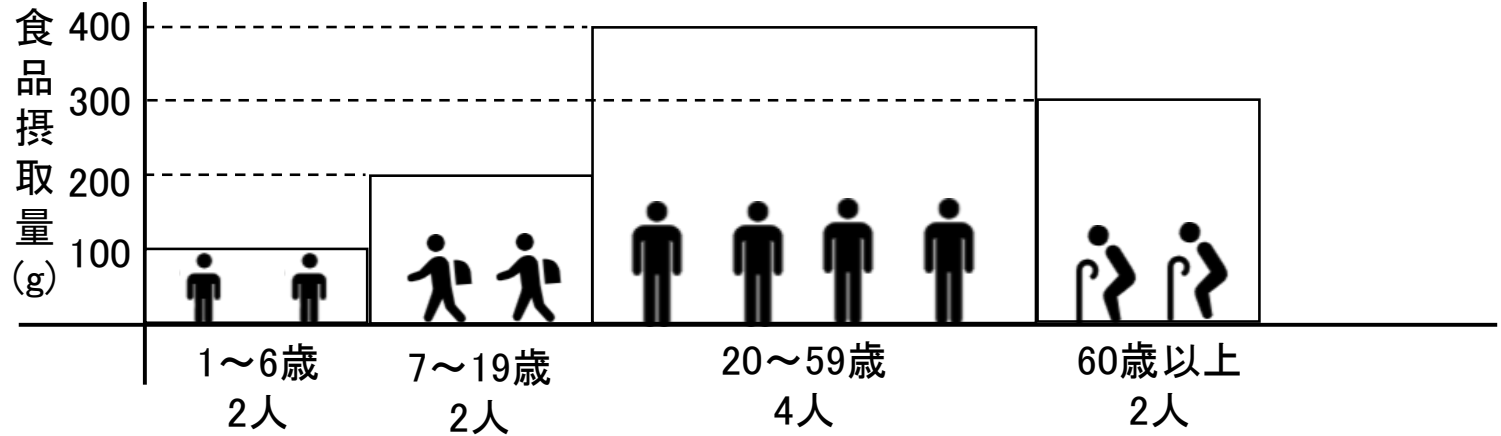
- 調査法、調査者、被調査者
 - 調査法の特徴
 - ✓ 食事記録法、24時間思い出し法、食物摂取頻度調査法の特徴
 - 調査者のスキル
 - ✓ 聞き取りが必要な調査法ではトレーニングが必要。
 - 被調査者の回答特性
 - ✓ 食事調査のある日は、いつもよりも簡素な(あるいは豪華な)食事。
 - ✓ 肥満者ではエネルギー摂取を過小申告しやすい傾向。

- ✓ 集団代表性のある標本抽出について
- ✓ 標本調査の誤差について
- ✓ 測定誤差について
-  ✓ 人口構成による重み付け
- ✓ 習慣的な摂取量の分布推定
- ✓ 体重あたり摂取量の分布について

人口構成を反映した食品摂取量

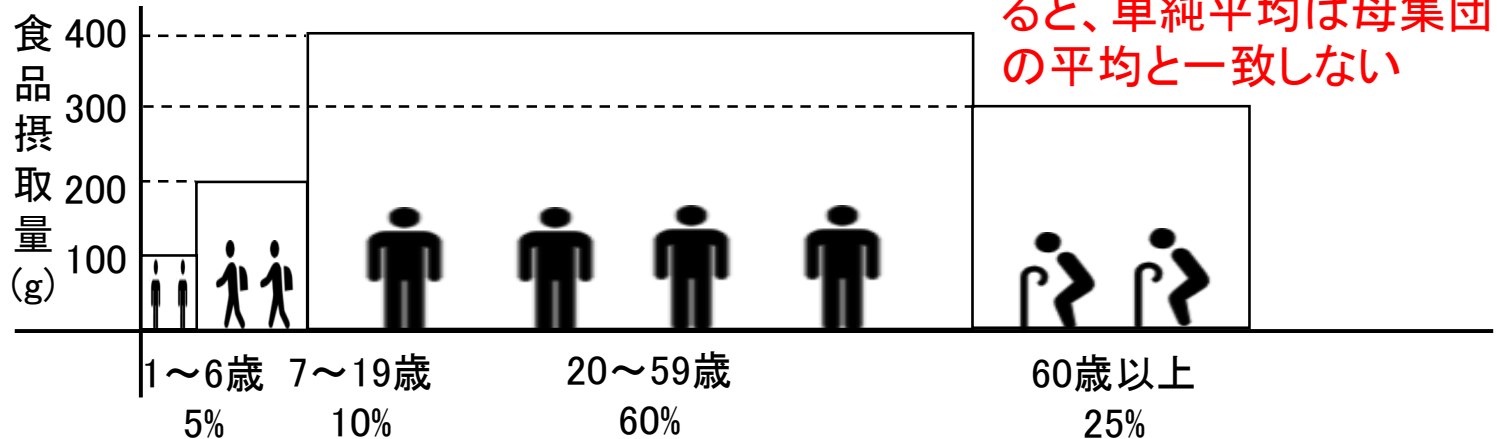
年齢によって食品摂取量が異なる

<調査集団の人口構成> (標本)



平均値: $(100+100+200+200+400+400+400+400+300+300) \div (2+2+4+2) = 280\text{g}$

<人口構成を反映> (母集団)



母集団と年齢構成が異なると、単純平均は母集団の平均と一致しない

平均値: $100 \times \frac{0.05}{2} + 100 \times \frac{0.05}{2} + 200 \times \frac{0.10}{2} + 200 \times \frac{0.10}{2} + 400 \times \frac{0.60}{4} + 400 \times \frac{0.60}{4} + 400 \times \frac{0.60}{4} + 400 \times \frac{0.60}{4} + 300 \times \frac{0.25}{2} + 300 \times \frac{0.25}{2} = 340\text{g}$

母集団の平均値を推定するために、重み付け計算を行う

人口構成による重み付け

割合小さい
⇒重みを大きく

- 一般に、性・年齢によって食事(食品)摂取量は異なる。
- そのため、全国から無作為抽出した調査であっても、協力者の性・年齢構成が国民全体と異なっていると(性・年齢で協力率が異なるために起きやすい)、その単純平均・割合・分布は、国民全体からずれる可能性がある。
- 国民全体の性・年齢構成に合わせるために重み付けを行う必要がある。

(仮想データ)	国民全体		調査協力者		野菜 摂取量 平均(g)	
	人口 (千人)	割合	人数	割合		
男性	20-29歳	5,934	5.9%	316	3.2%	231
	30-39歳	6,359	6.3%	387	3.9%	243
	40-49歳	8,090	8.0%	615	6.2%	238
	50-59歳	9,060	8.9%	827	8.3%	260
	60-69歳	7,215	7.1%	713	7.1%	272
	70-79歳	7,396	7.3%	787	7.9%	284
	80歳以上	4,720	4.7%	538	5.4%	280
	女性	20-29歳	5,666	5.6%	377	3.8%
30-39歳		6,107	6.0%	464	4.6%	220
40-49歳		7,822	7.7%	743	7.4%	222
50-59歳		8,886	8.8%	1,013	10.1%	238
60-69歳		7,435	7.3%	918	9.2%	263
70-79歳		8,593	8.5%	1,143	11.4%	289
80歳以上		8,125	8.0%	1,158	11.6%	266
男女		総数	101,408	100.0%	10,000	100.0%

単純平均 256 g
重み付け平均 252 g

割合大きい
⇒重みを小さく

- ✓ 集団代表性のある標本抽出について
- ✓ 標本調査の誤差について
- ✓ 測定誤差について
- ✓ 人口構成による重み付け
- ✓ 習慣的な摂取量の分布推定
- ✓ 体重あたり摂取量の分布について

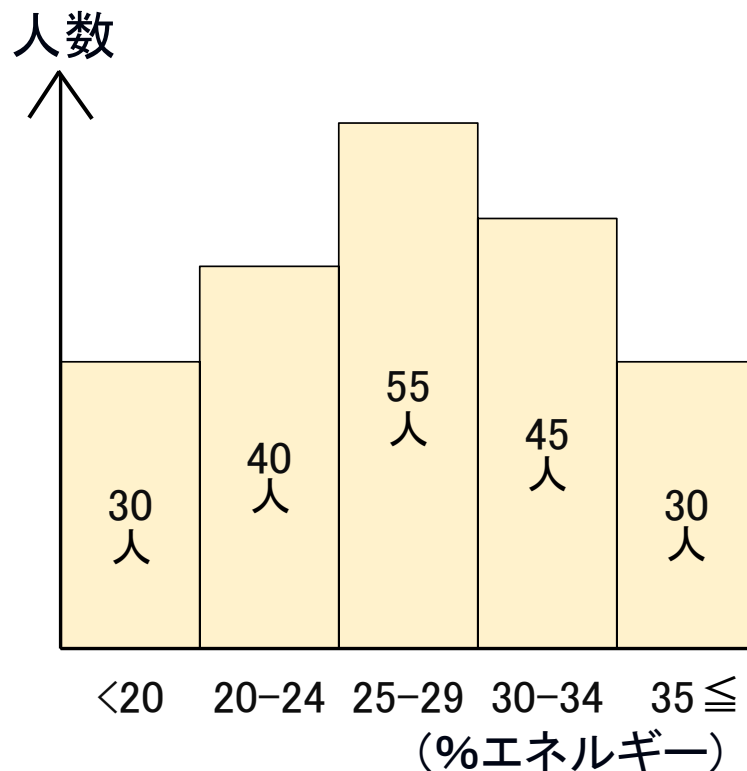


例示として栄養素を示しますが、他の
化学物質についても考え方は同様

習慣的な摂取量の分布は1日調査では測れない

1日間食事記録調査を成人200人に行いました。脂肪エネルギー比率（総エネルギーに占める総脂質エネルギーの割合）の分布は右のグラフの通りです。

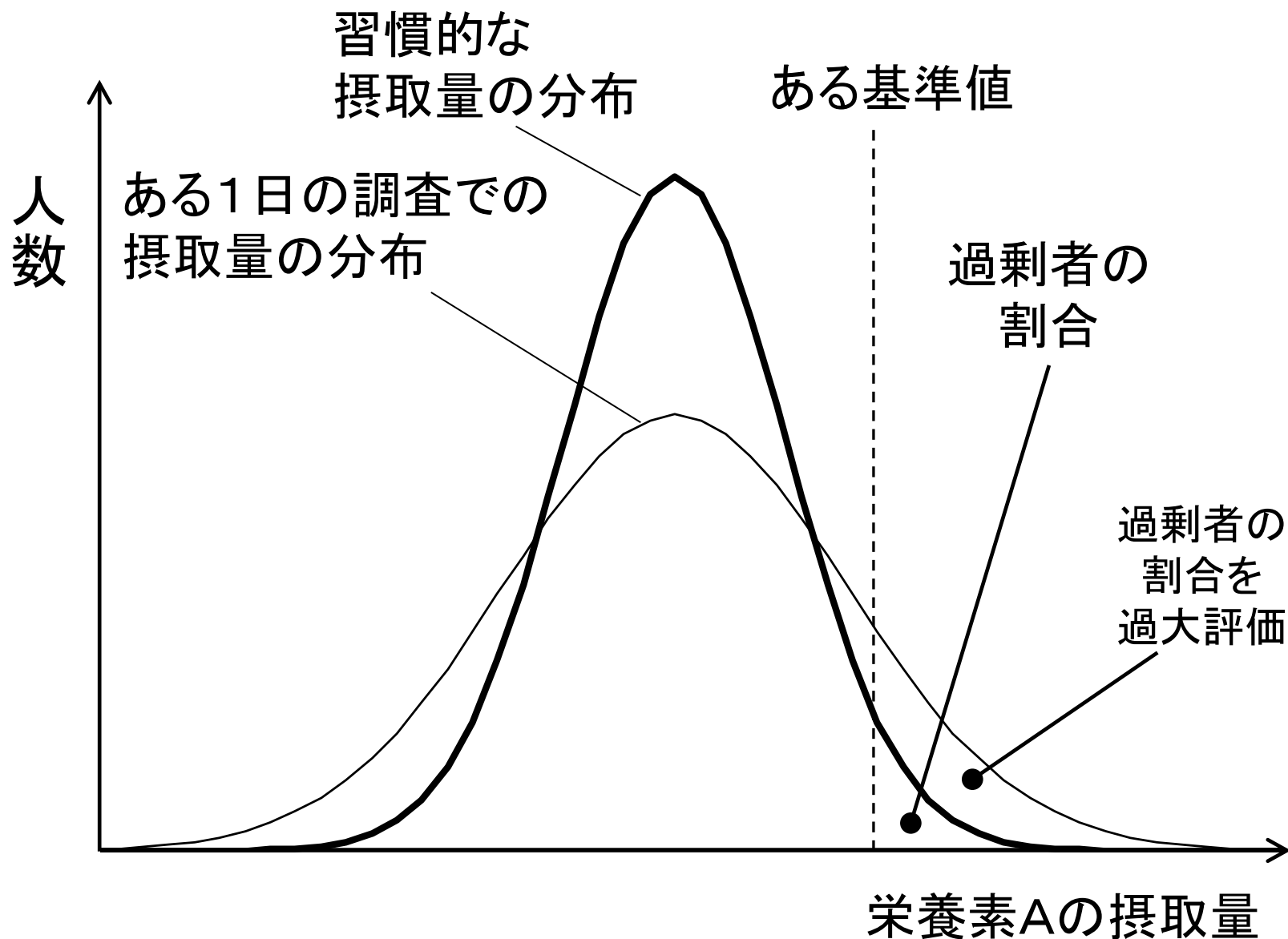
習慣的な脂肪エネルギー比率が35%以上の方は



- (1) 多分30人いる。
- (2) 多分30人より多い。
- (3) 多分30人より少ない。

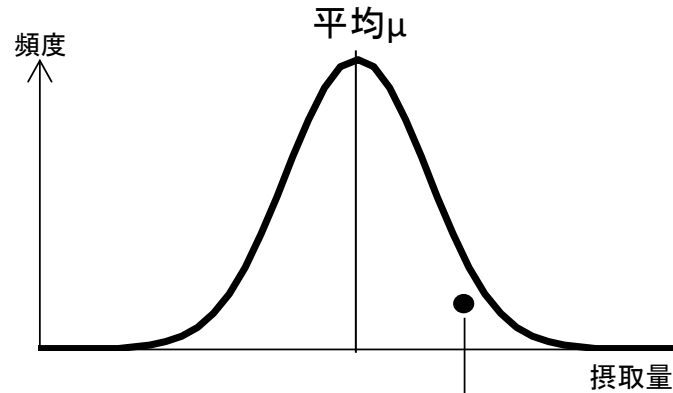
注: 食事摂取基準の目標量の上限は30%ですが、極端な場合として35%で例示しました。

集団における栄養素Aの習慣的な摂取量と1日調査での摂取量の分布の違い
(仮想的な模式図)。

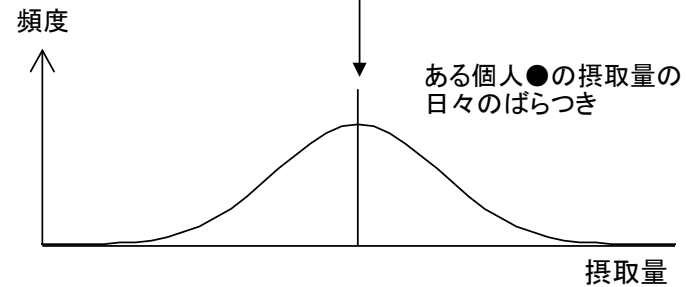


習慣的な摂取量の分布とある1日の調査による摂取量の分布との関係 (仮想的な模式図)。

習慣的な摂取量
の分布
(分散 σ_b^2)

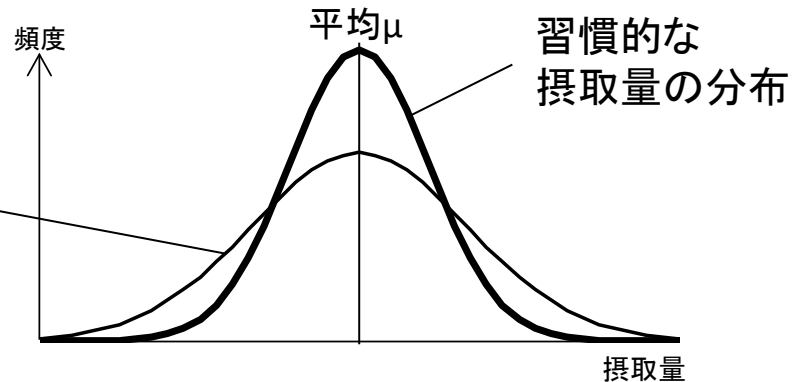


日々のばらつき
(分散 σ_w^2)



+

ある1日の調査に
よる摂取量の分布
(分散 $\sigma^2 = \sigma_b^2 + \sigma_w^2$)



=

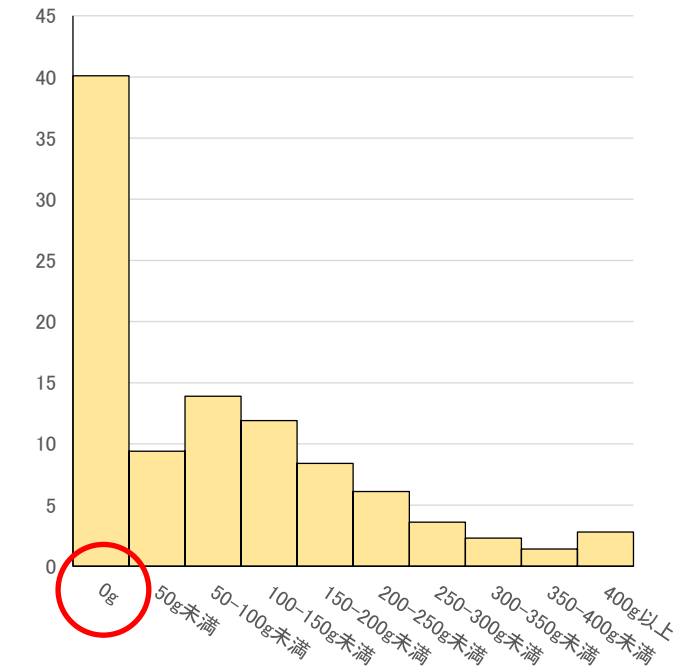
栄養素の習慣的な摂取量の分布を調べる方法

- 多数日の食事調査を行って、その平均値をその人の摂取量とする
 - 栄養素によっては膨大な日数が必要であり、現実的でない。
- 統計学的手法で推定する(複数日調査が必要)
 - National Research Council (NRC)法 (*National Research Council, 1986*)
 - 分散分析(ANOVA)を応用した比較的簡単で基本的な方法。
 - 正規分布でない栄養素を扱いにくく、あまり実用的でない。
 - Best-Power (BP)法 (*Nusser et al., 1996*)
 - 正規分布と対数正規分布のみならず、高値側に裾が長い様々な程度に歪んだ分布を扱うことができる。
 - 習慣的な摂取量＝長期間の摂取量の算術平均とみなすことができ、より現実のデータを扱いやすい。
 - ISU法の簡易版という位置づけで、性能はあまり劣らない。
 - 日本語ソフト(HabitDist※)が提供されている。
 - Iowa State University (ISU)法 (*Nusser et al., 1996*)
 - BP法に比べて、より正確に正規近似が可能。
 - 個人内分散が個人ごとに異なる現実に近い状況を許容。
 - 英語版ソフト(PC-SIDE)が提供されている。

食品の習慣的な摂取量の分布を調べる方法

- 前記の統計手法は、正規分布(または変換して正規分布)を仮定しているため、**摂取量0が多い食品には使えない。**
- 対応した統計学的手法(※1)
 - Iowa State University for estimating the distribution of foods (ISUF). (*Nusser et al. 1996*)
 - National Cancer Institute (NCI) method. (*Janet A Tooze et al. 2006*)
 - Multiple Source Method (MSM). (*Haubrock et al. 2011*)
 - Statistical Program to Assess Dietary Exposure (SPADE). (*Dekkers et al. 2014*)
 - 性能に大きな違いはないという報告がある(※2)。
 - いずれも計算ソフトが提供されている(※1)。

1日調査での果物摂取量の分布
(令和5年国民健康・栄養調査)男女総数

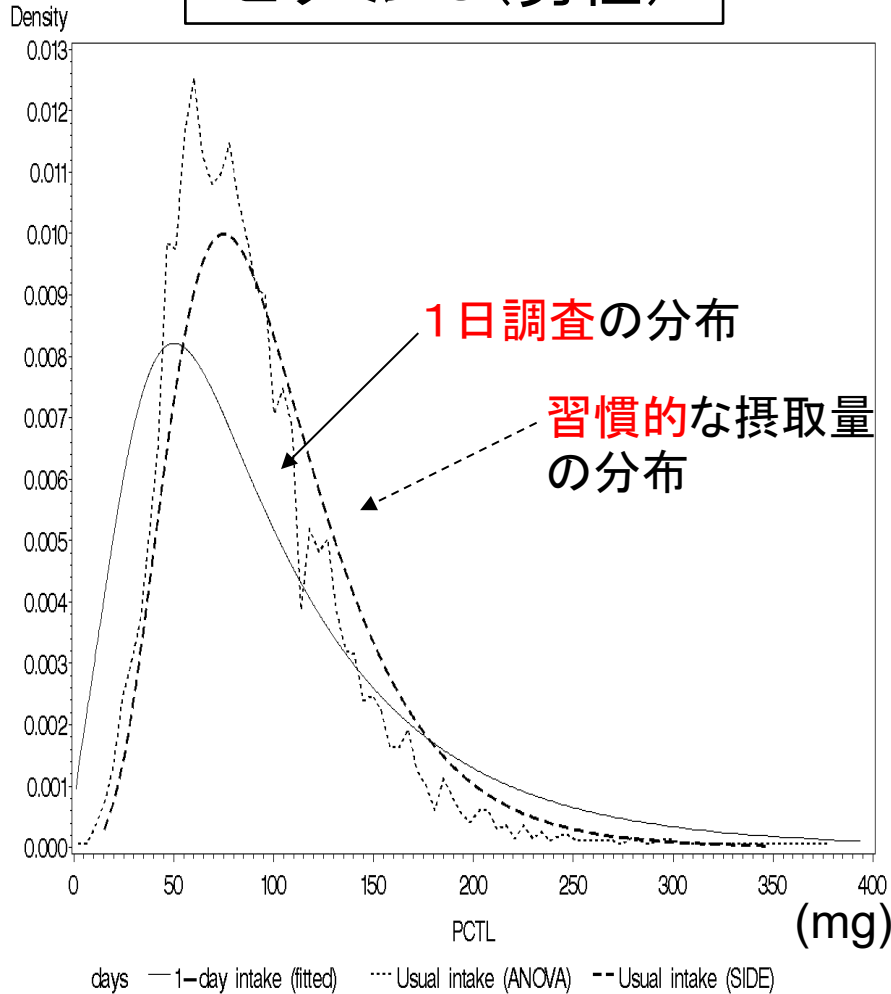


※1 Hilko van der Voet et al. Statistical modelling of usual intake. <https://doi.org/10.2903/sp.efsa.2010.EN-86>

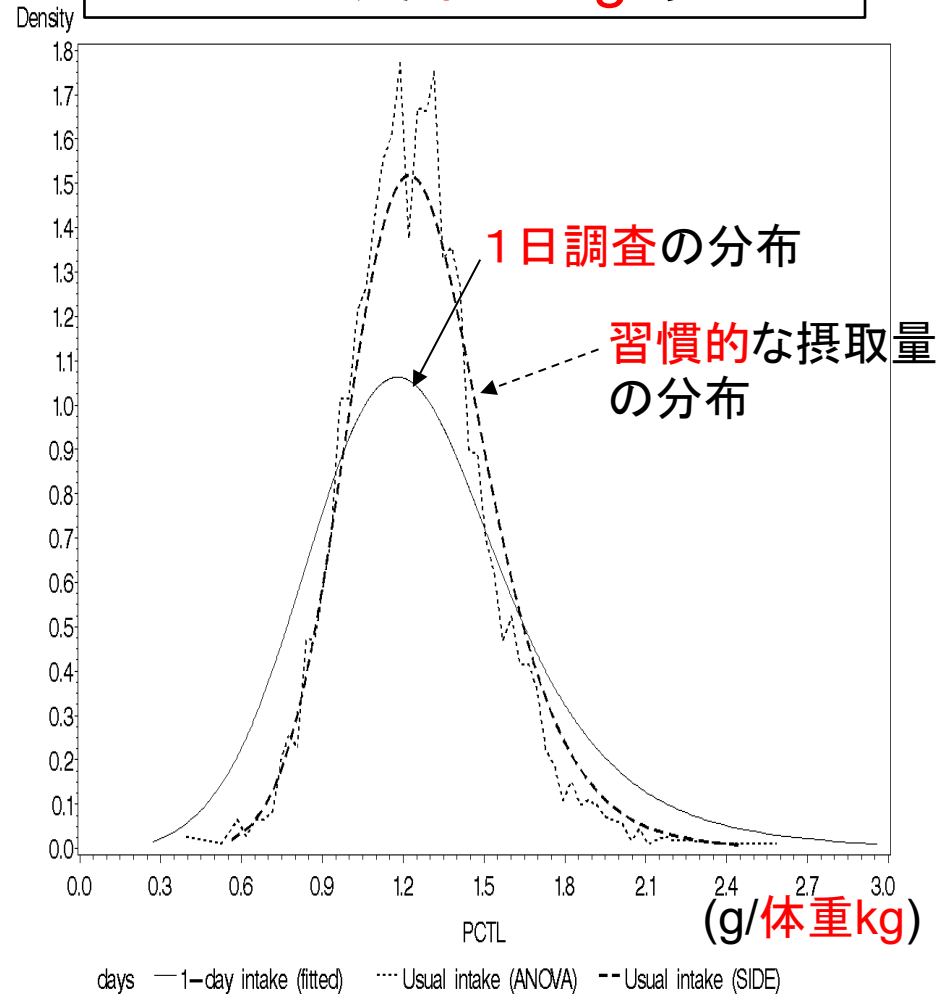
※2 Greice HC Laureano et al. Comparison of the ISU, NCI, MSM, and SPADE Methods for Estimating Usual Intake: A Simulation Study of Nutrients Consumed Daily. *Nutrients*. 2016;8(3):166. doi: 10.3390/nu8030166.

平成14年国民栄養調査における、1日摂取量と推定習慣的摂取量の分布

ビタミンC(男性)



たんぱく質/体重kg(男性)



- ✓ 集団代表性のある標本抽出について
- ✓ 標本調査の誤差について
- ✓ 測定誤差について
- ✓ 人口構成による重み付け
- ✓ 習慣的な摂取量の分布推定
- ✓ 体重あたり摂取量の分布について



例示として栄養素を示しますが、他の
化学物質についても考え方は同様

体重あたり摂取量の分布について

A. 正しい計算方法

- 個人ごとに「摂取量÷体重」を計算してその分布を調べる。
 - 個人ごとのデータが必要。

B. 「摂取量の平均値」÷「体重の平均値」ではだめか？

- 集計データだけで計算できるが、Aとは異なる。
- 例)

➤ たんぱく質 平均 80g 標準偏差 26g

➤ 体重 平均 67kg 標準偏差 12kg

➤ 両者が正規分布、相関係数=0.1の場合

✓ 「摂取量の平均値」÷「体重の平均値」= $80 \div 67 = 1.194$ g/体重kg

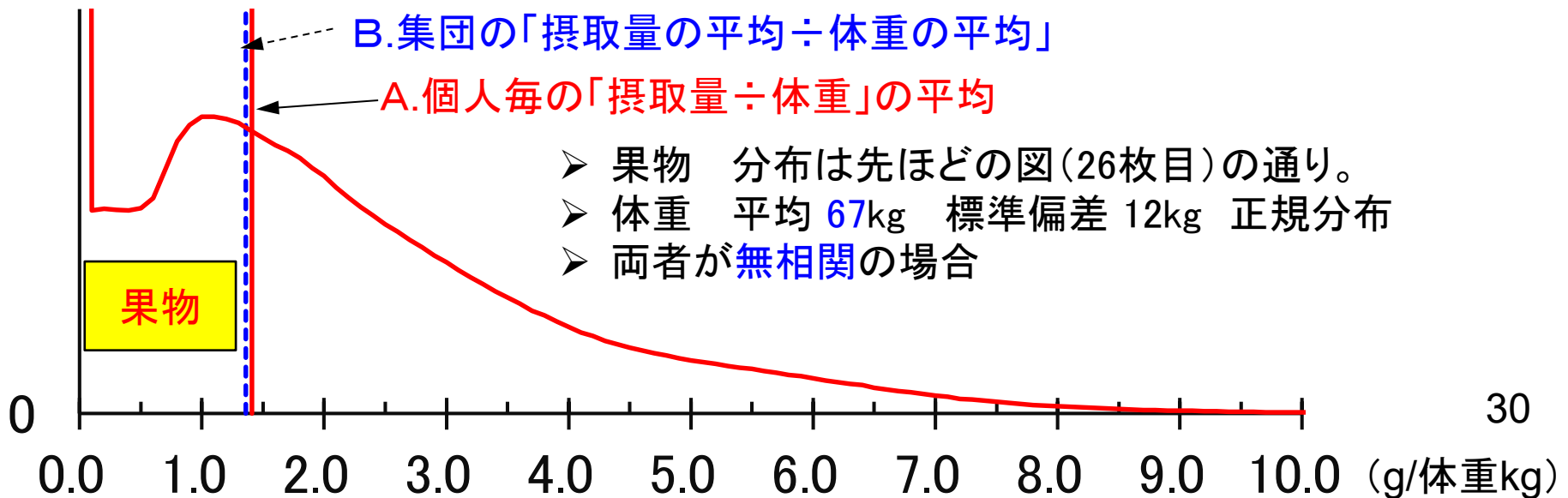
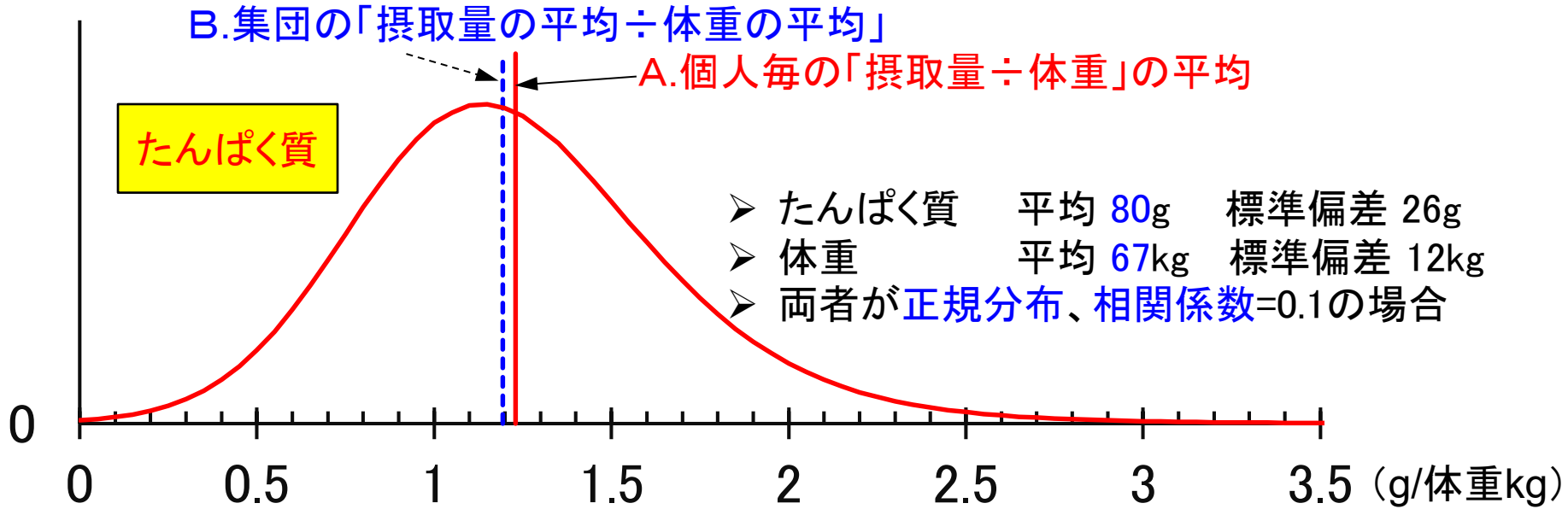
✓ 「摂取量÷体重」の平均値±標準偏差 = 1.229 ± 0.457 g/体重kg

➤ ずれの大きさは、標準偏差の約8% ($[1.229 - 1.194] \div 0.457$)。

➤ どの程度ずれるかは、分布型と相関の強さにもよる。

体重あたり摂取量の分布について(続き)

いずれもシミュレーションによる推定



まとめ

ある食品の習慣的な摂取量の分布を調べるために

- 母集団からの無作為抽出による標本調査
- 十分なサンプルサイズと協力率
- 測定誤差を減らす工夫
- 母集団と標本の性・年齢構成の補正
- 複数日調査データに統計学的手法を適用
- 体重の扱いに注意