

# 自然科学の歩き方

第2回目

# モデル

- ★ 「自然を理解する」と言っても、神ならぬ身である人間には全てを完全に理解することは（おそらく）不可能
- ★ 人間に理解できる形で自然を記述する必要
- ★ 現象を大雑把に理解する必要がある
  - ★ どれくらい大雑把でいいかというのは、真面目に誤差の議論をした上で決定される
  - ★ 実験誤差（統計誤差と系統誤差）と理論誤差（ある意味では系統誤差の一部）
- ★ そのために「モデル（模型）」を考える

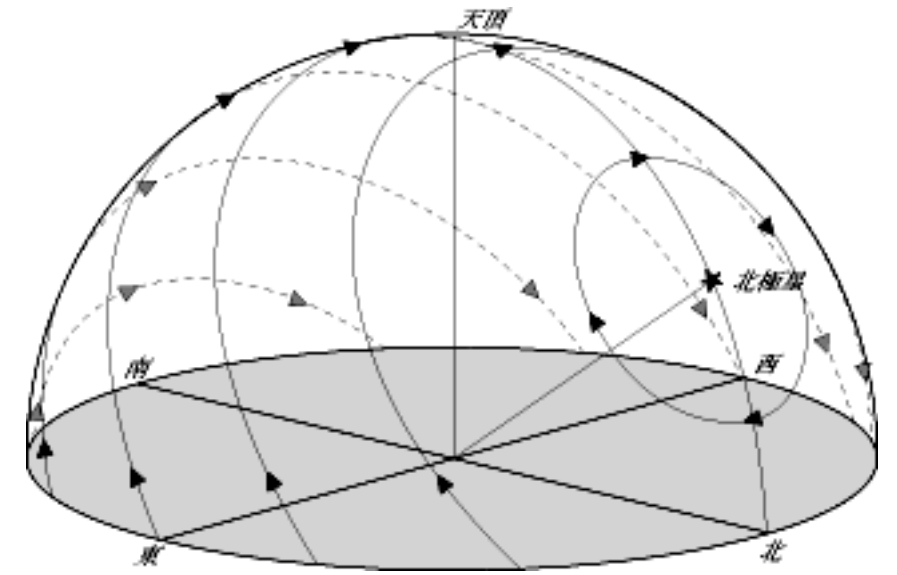
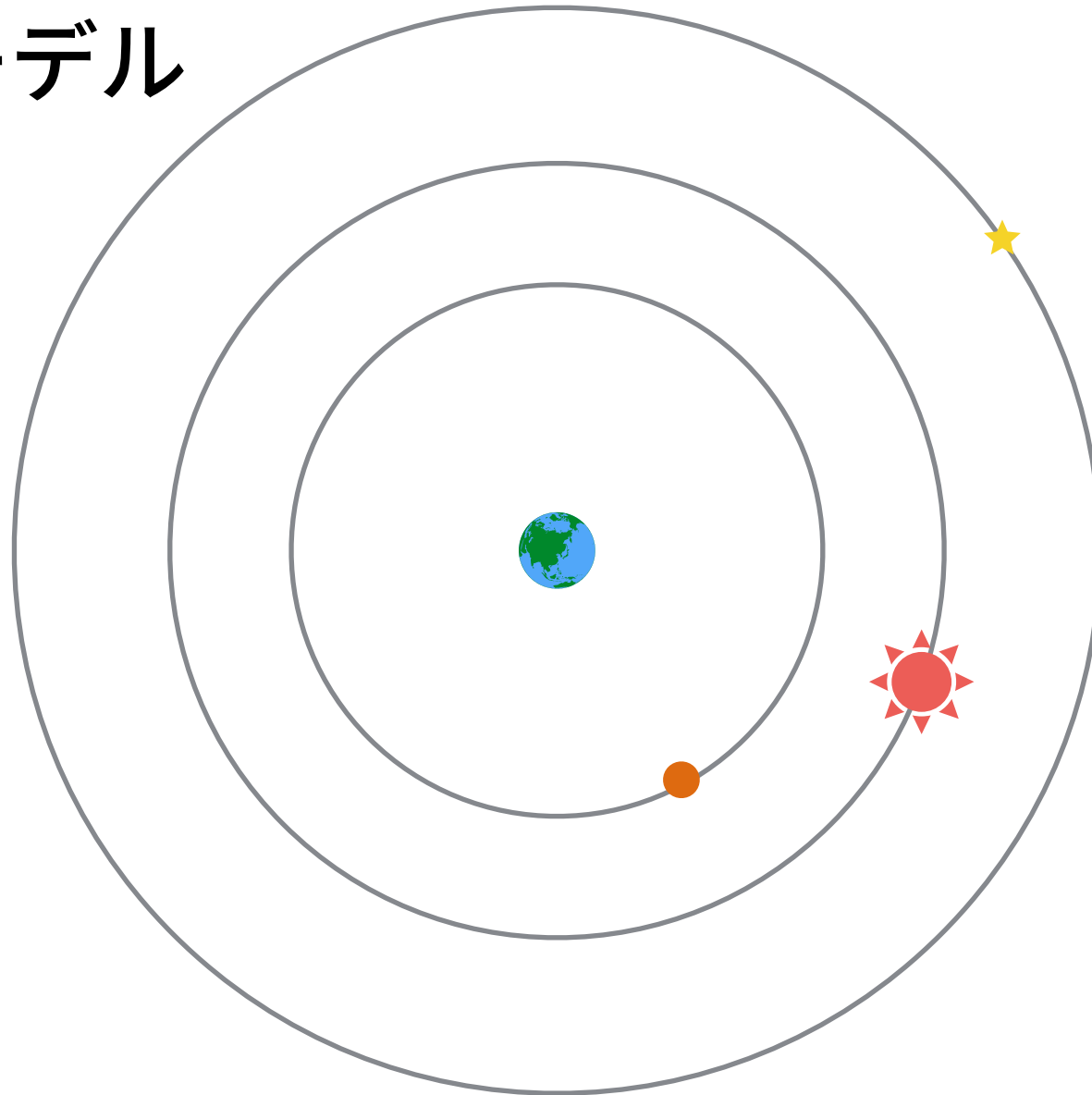
# モデルの例

- ★ 全ての物理学（自然科学）の法則は「モデル」である
- ★ 言い換えれば、「自然科学は近似の学問」である
- ★ 「良いモデル」を作るには、枝葉末節を取っ払って、現象の本質を見極めることが重要になる
- ★ 質点の運動
- ★ 自由落下
- ★ ...

# モデルの例

太陽系の惑星運動

単純なモデル



「AstroArts」のサイトより

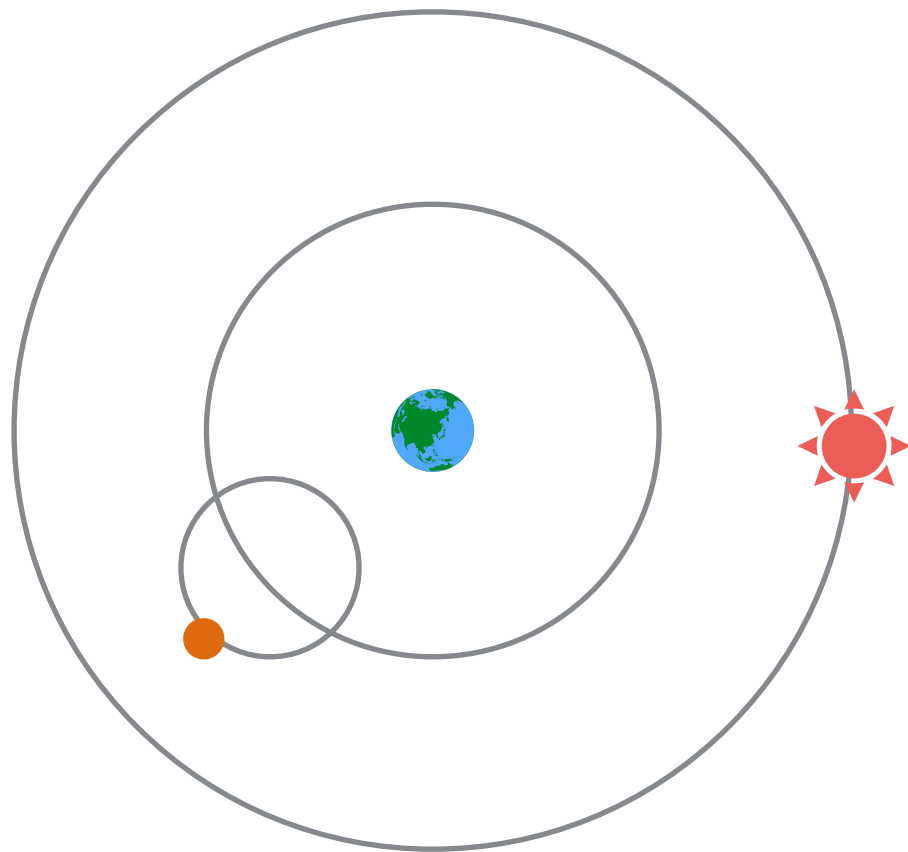
日周運動は説明できる。

# モデルの例

## 太陽系の惑星運動

単純な天動説だと，惑星の運動を説明できない→改良

## 周転円モデル

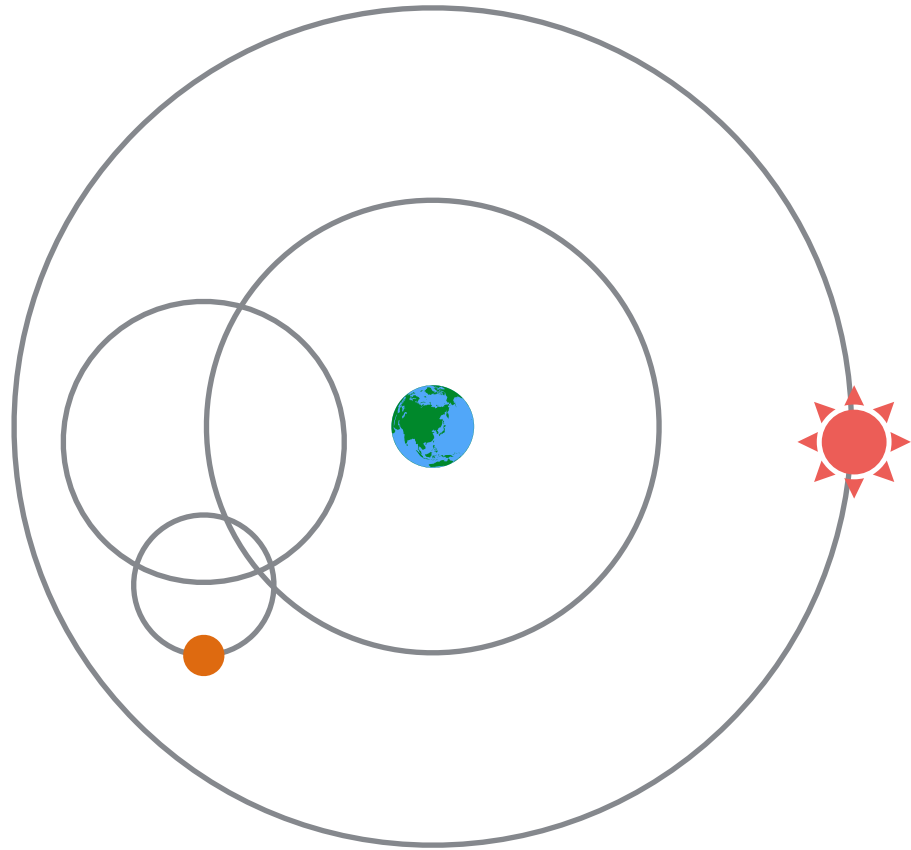


<http://www.phys.u-ryukyu.ac.jp/~maeno>

# モデルの例

太陽系の惑星運動

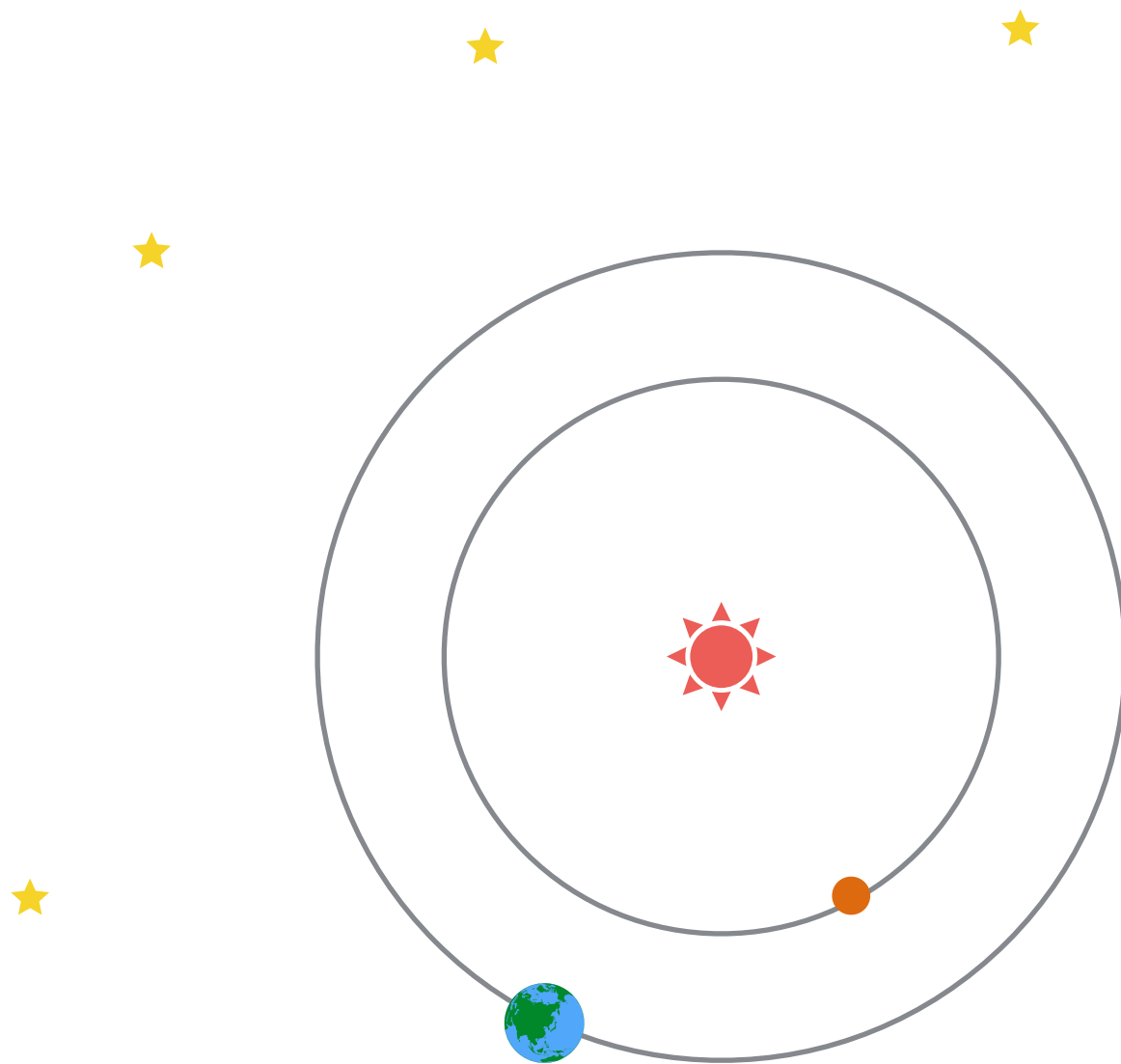
観測が精密化→モデルが複雑化



# モデルの例

## 太陽系の惑星運動

### 単純な模型としてのコペルニクスの地動説



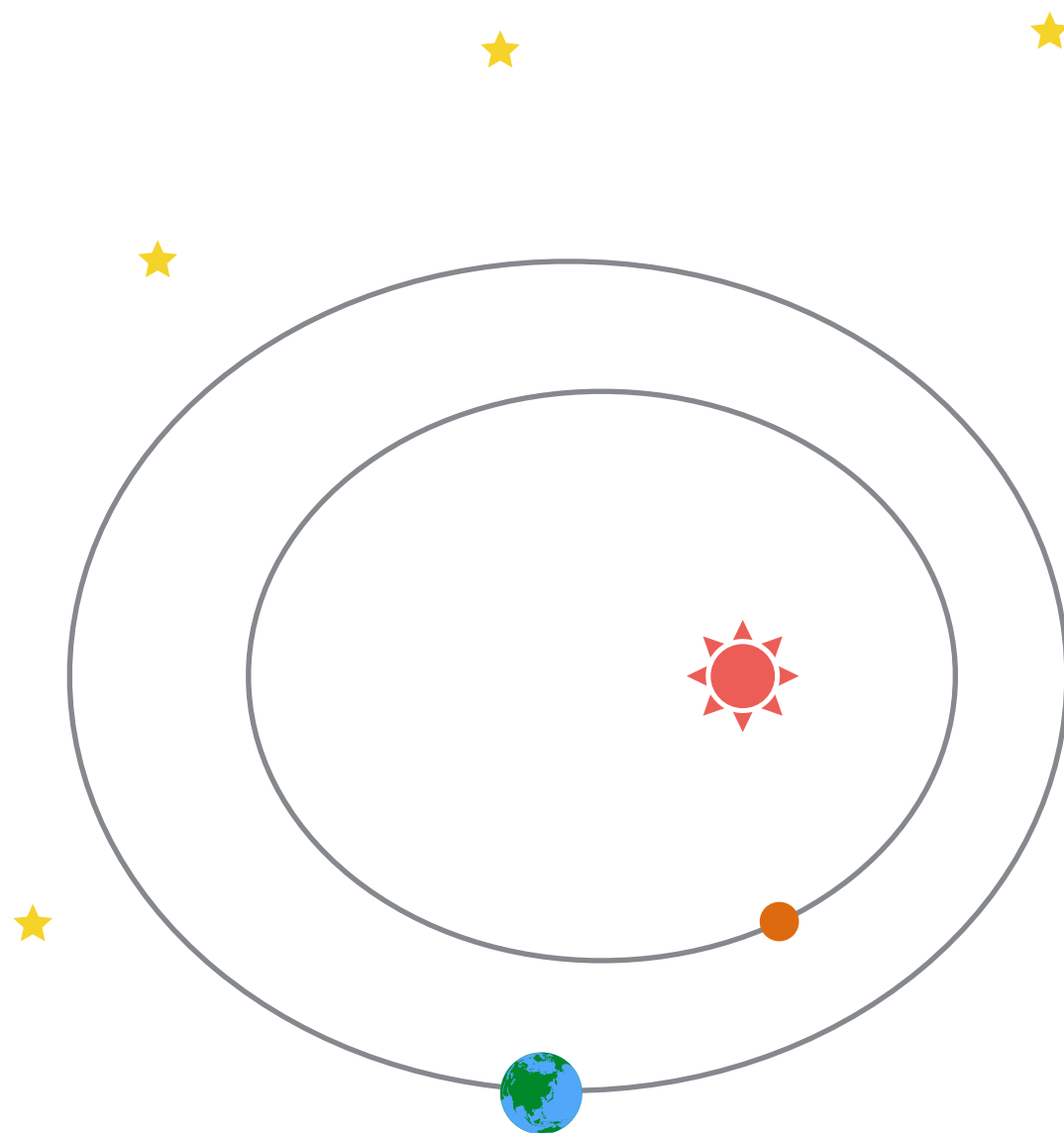
いくつかの点で観測と不一致

- 周転円模型より惑星データとの不一致が大きい
- 年周視差が観測されない

# モデルの例

太陽系の惑星運動

ケプラー模型



- 惑星運動データを非常に良い精度で再現
- 年周視差の問題→惑星以外の星は非常に遠くにある



# 演繹法と帰納法

★ モデルの作り方にも色々な方法がある

★ 第一原理に基づいてモデルを作る (演繹法)

★ 例：相対性理論

★ 現象論的方法 (帰納法的)

★ 実験・観測に基づいて、モデルを作る

# データの推定

★ 測定されていないデータは知りようがない

★ 測定データを元に内挿・外挿する必要あり

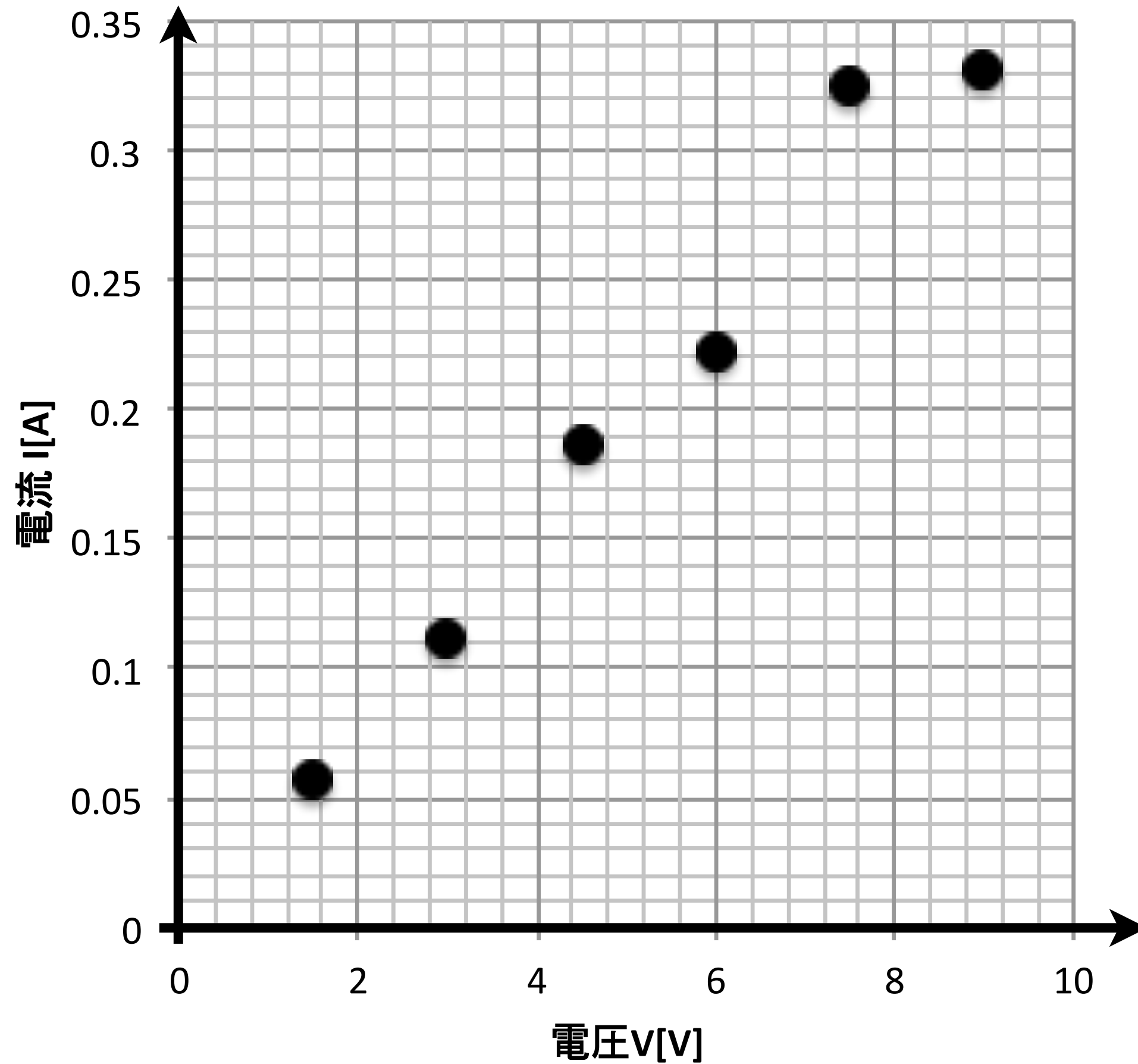
★ 例：2.5Vの電圧に対する電流値はいくらか？

V[V]	1.50	3.00	4.50	6.00	7.50	9.00
I[A]	$5.64 \times 10^{-2}$	$1.12 \times 10^{-1}$	$1.86 \times 10^{-1}$	$2.22 \times 10^{-1}$	$3.25 \times 10^{-1}$	$3.32 \times 10^{-1}$

おそらく0.0564Vと0.112Vの間

★ 結果の見当をつけることは重要なステップ

★ ただし、実際のところは測定してみるまでわからない

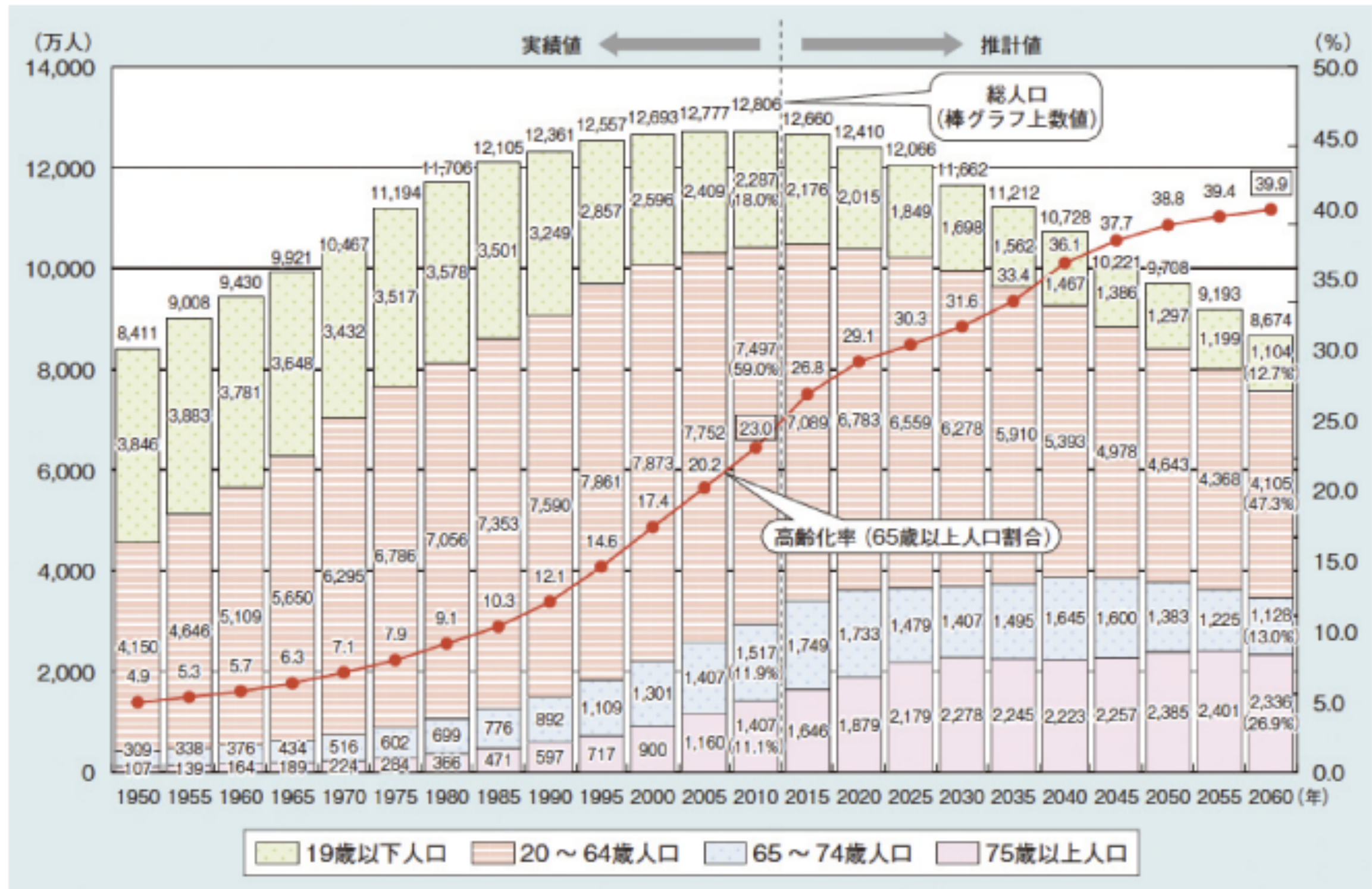


# データの推定をするために

## ★ 数学を利用する

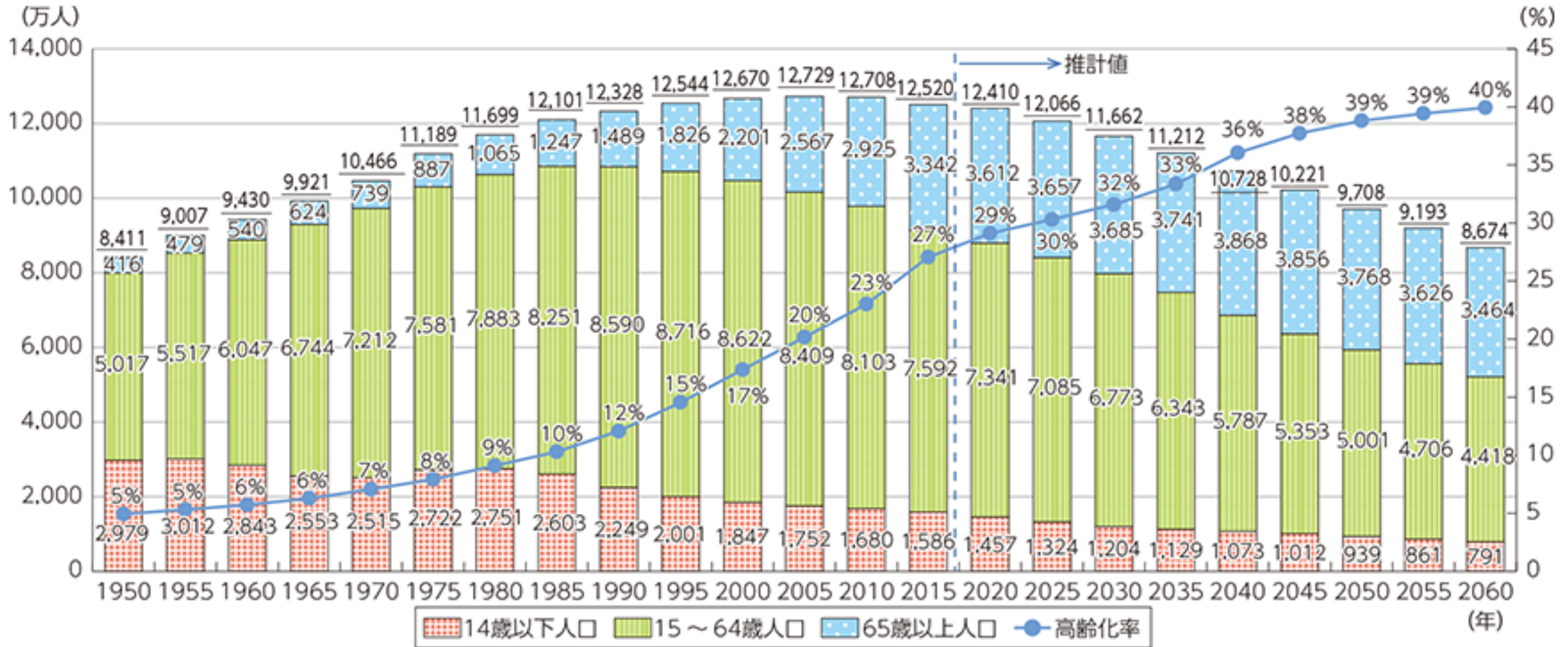
- ★ 関数を数式で表しておけば、（定義域内の）任意のインプットに対してアウトプットを得ることができる
- ★ 測定されたデータを，数式を使って「近似的に」表してみる
- ★ こうしておけば，測定値がないところでも測定値を**予言**(≠ 預言)することができる
- ★ どうやって適切な数式を求めていくか？

# 実測値，モデル，予言





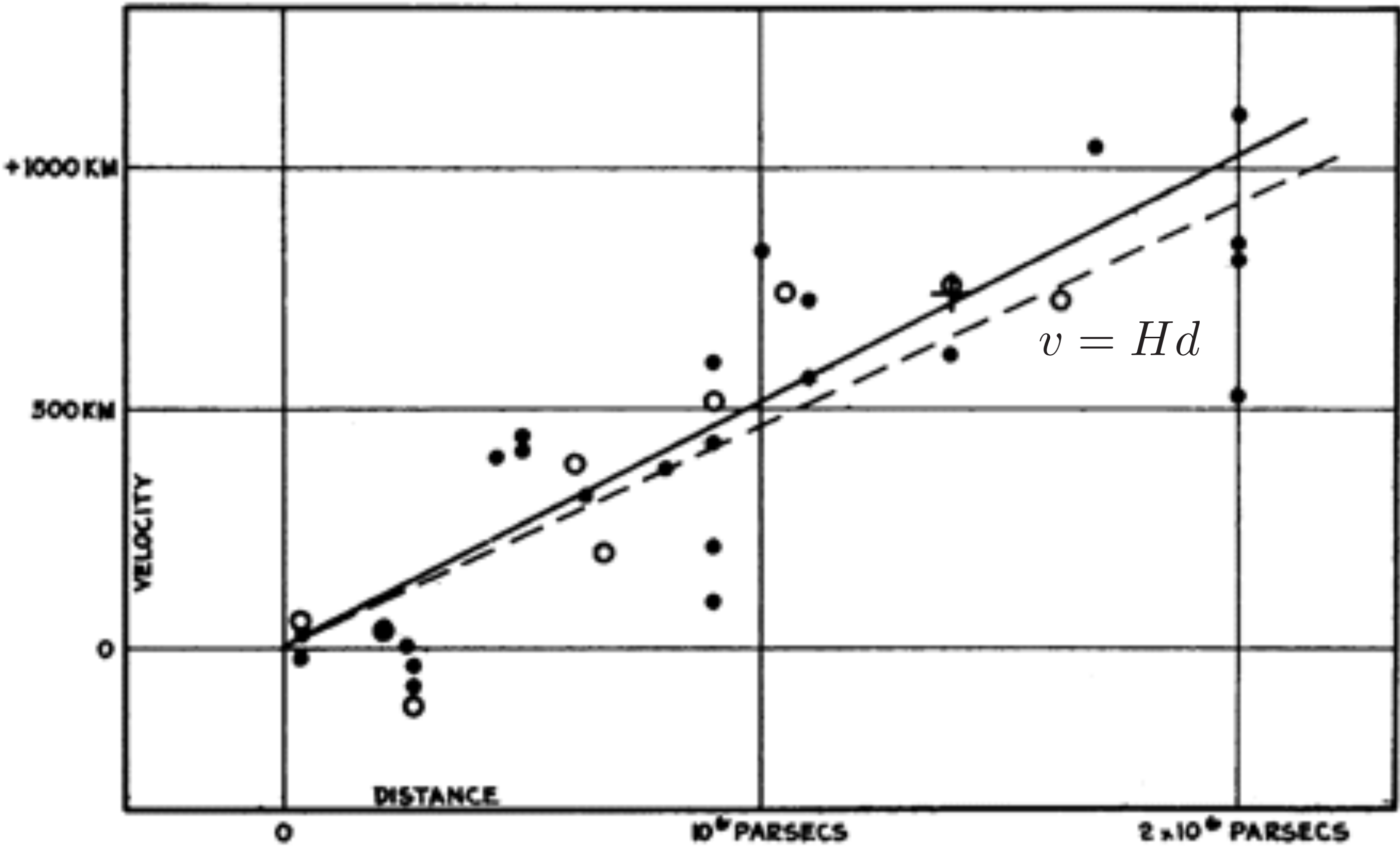
# 実測値，モデル，予言



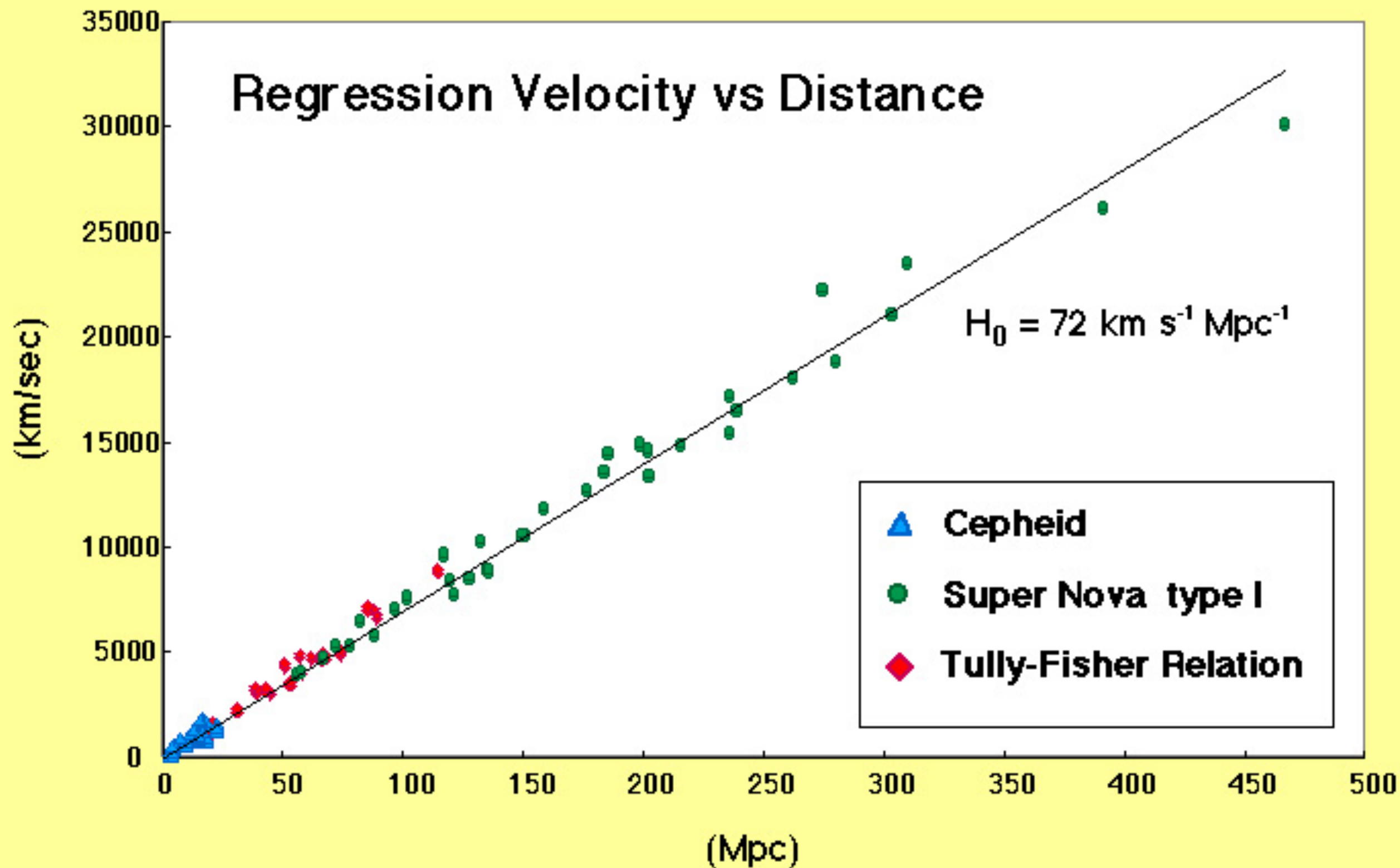
平成28年度情報通信白書

# モデル作りに必要なこと

- ★ 測定データを信用しすぎない
  - ★ 測定や実験には必ず**誤差**が入り込む
  - ★ 誤差評価のないデータは全く意味がない
  - ★ 誤差の扱いについては，実験関係の授業で詳しく学ぶ
- ★ ある程度単純な数式で表現する
  - ★ 誤差を考慮すると，中心値を完璧にあわせることに意味はない。
  - ★ 単純=モデルパラメータが少ない







データ出典： Freedman, W. L. et al (2001) *Astrophysical Journal*, **553**, 47-72

# オッカムのカミソリ



- ★ 「何かを説明するのに，必要以上に多くのことを仮定するべきではない」
- ★ 同じ事柄を説明できる複数の仮説があった場合，より少ない仮定で説明できる仮説がよい仮説である。
- ★ 自然科学の場合，より高い精度の，よりシンプルな理論の構築を目指すべきという思想

# モデルパラメータの話

★ モデルパラメータとは何か？

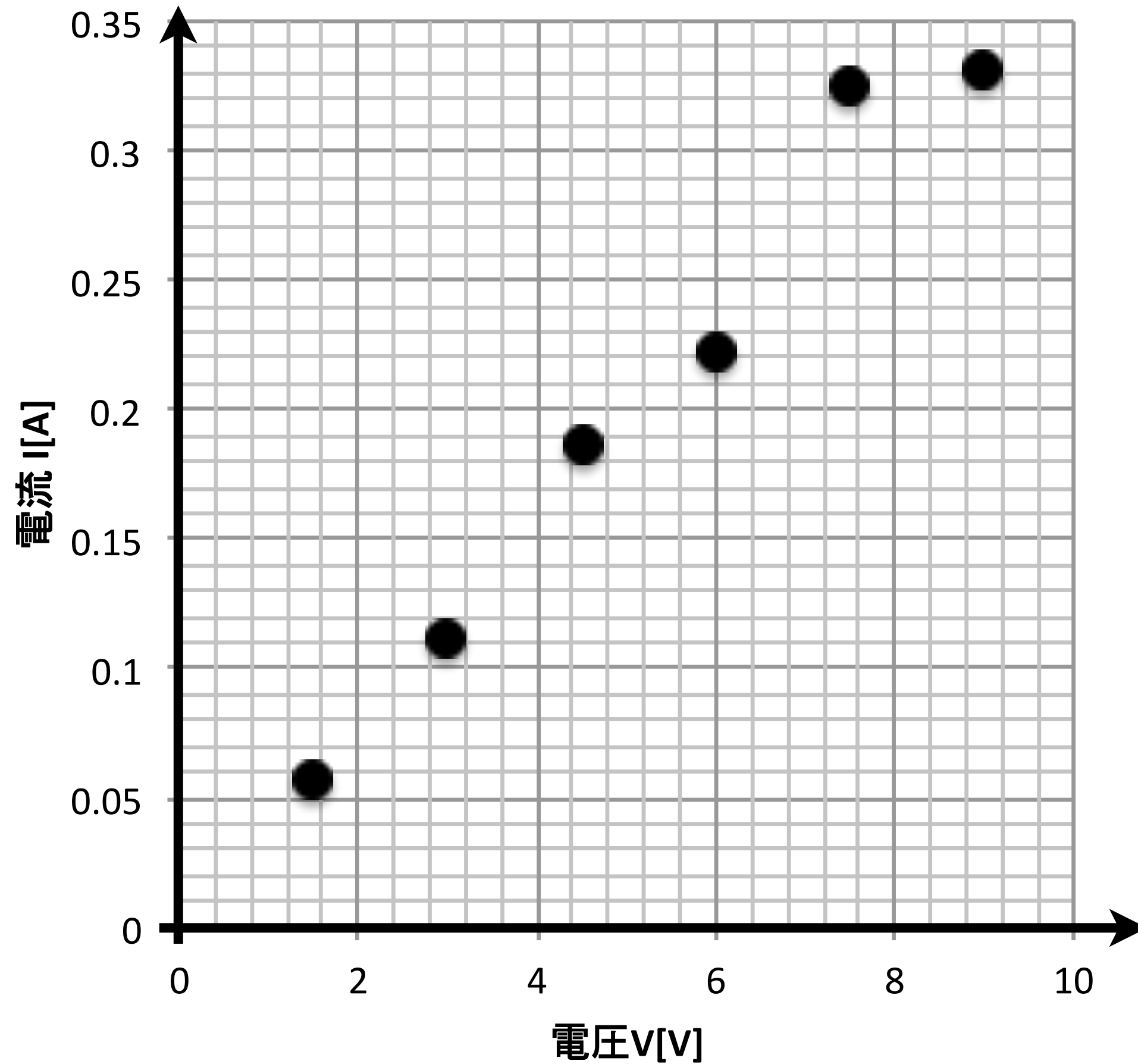
★ モデルに登場するパラメータのこと

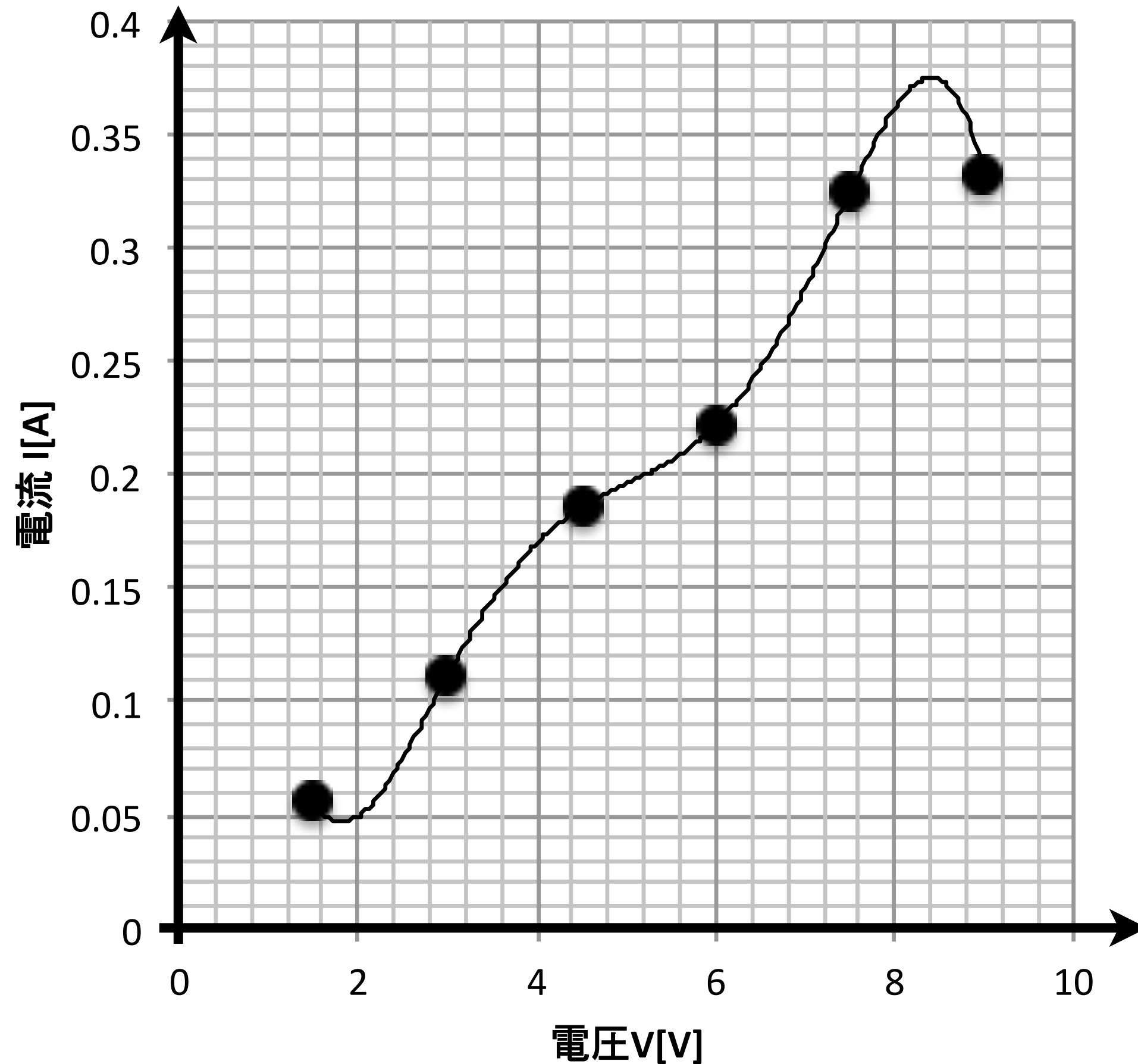
★ 例えば，電流 $I$ を電圧 $V$ の5次関数で表すモデルの場合

$$I = c_0 + c_1V + c_2V^2 + c_3V^3 + c_4V^4 + c_5V^5$$

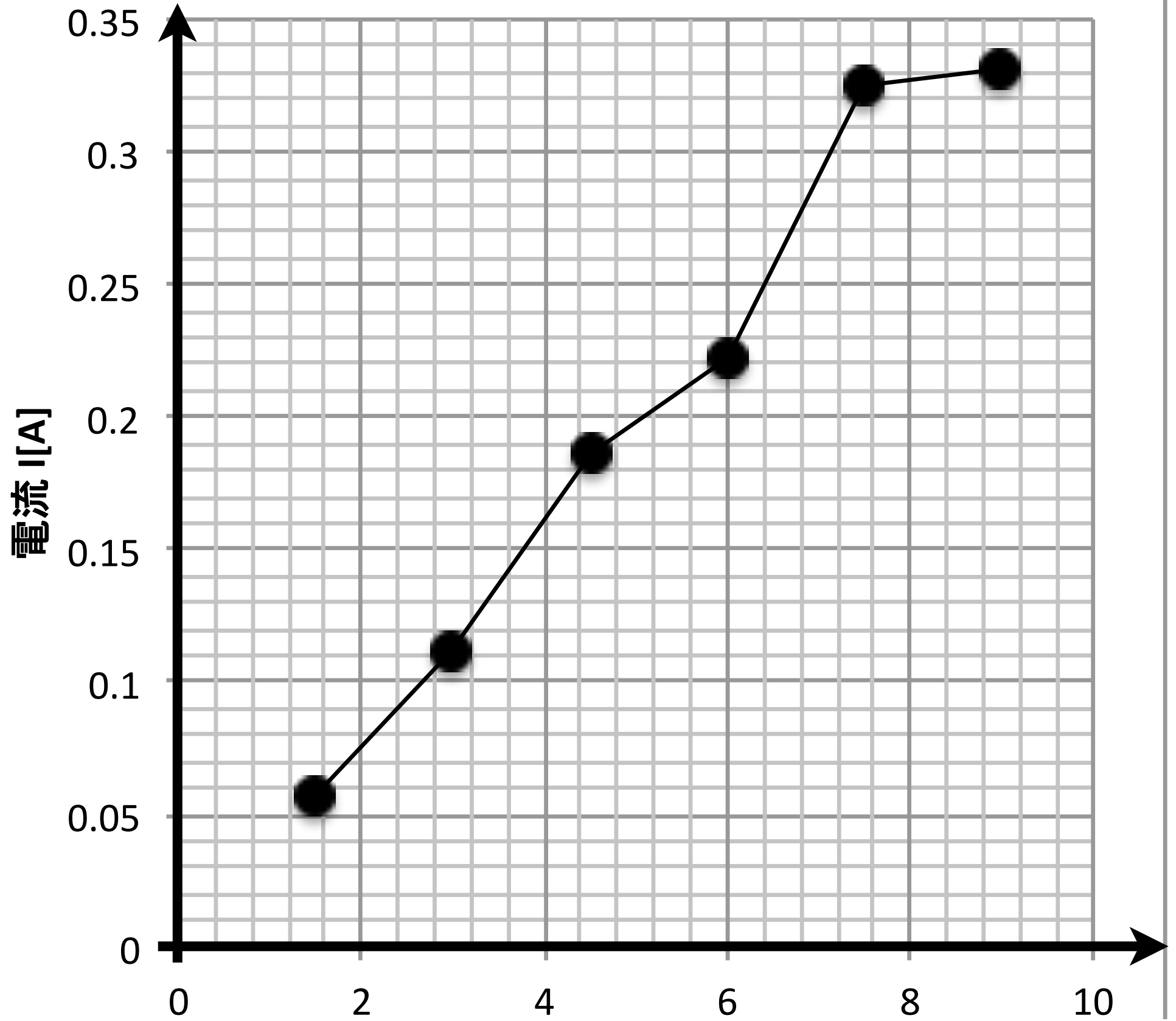
★ 関与するパラメータは6個。これらの値を全て決めるのに6個のインプットが必要。

★ パラメータが少ないモデルほど予言能力は高くなる





$$I = 0.667 - 0.928V + 0.492V^2 - 0.113V^3 + 0.0119V^4 - 0.00047V^5$$



これはさらに悪い  
(10パラメータ)

# もっとシンプルに

★ いっそのこと，比例関係を仮定してみてもどうか？

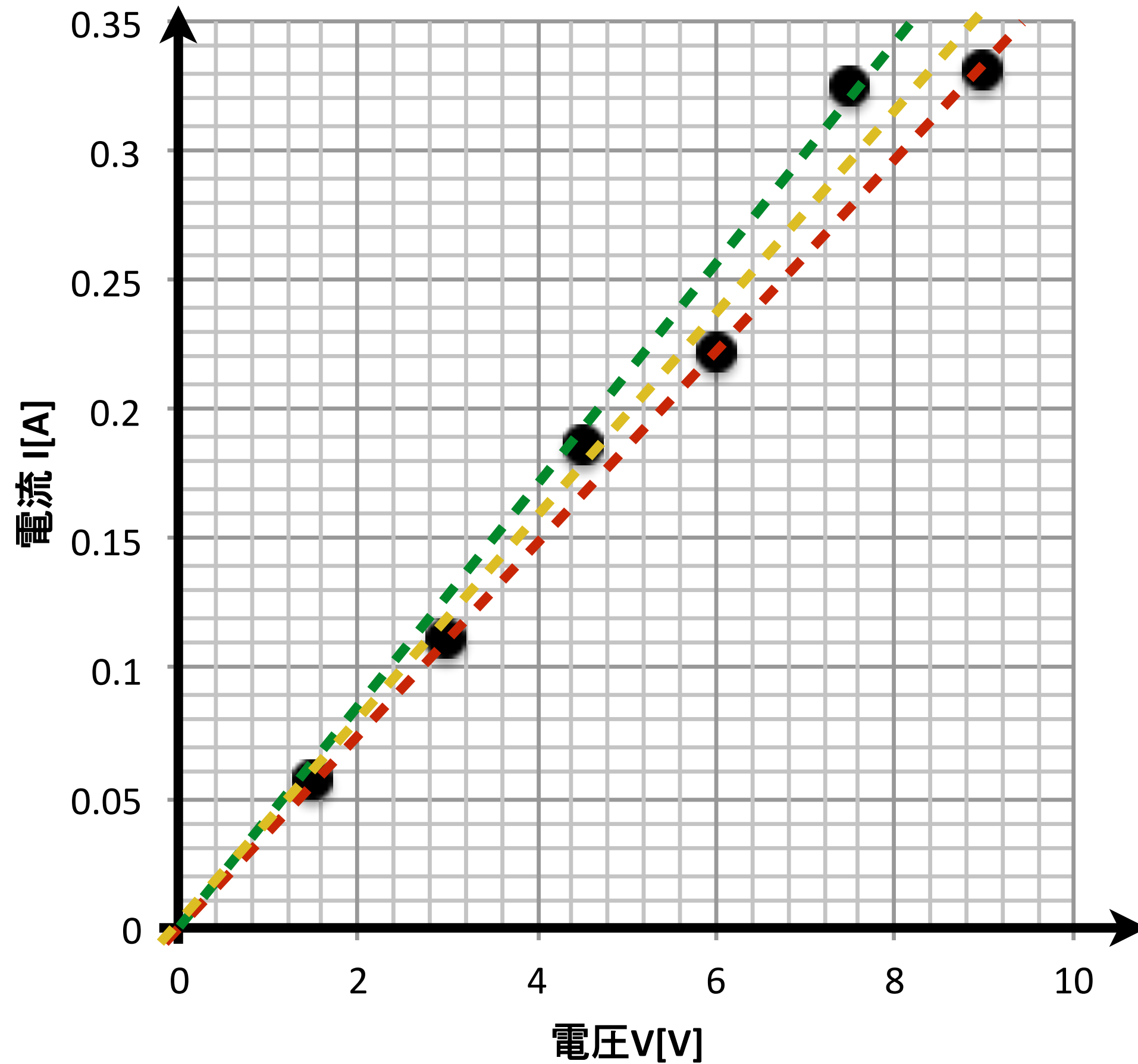
$$I = aV$$

★ パラメータは1つだけ

★ 全ての測定値を完璧に再現できるわけではない

★ 測定値に誤差があることは思い出す必要あり

★ 傾きの値をどう設定するべきか？





# モデルに必要なこと

- ★ 全てのデータを「それなりの」精度で再現できること
- ★ モデルの予測がはっきりわかる単純さ
- ★ 必要最小限の拡張とパラダイムシフトを繰り返しつつ、モデルは洗練されていく

# モデルの評価

- ★ モデルの良し悪しをどう判断するか？
  - ★ 測定データをそれなりに再現できるか？
  - ★ 式として可能な限り単純であるか？
- ★ 「それなり」とは？
  - ★ 客観的な評価が必要
  - ★ さもないと喧嘩になる

# データの再現性を評価する

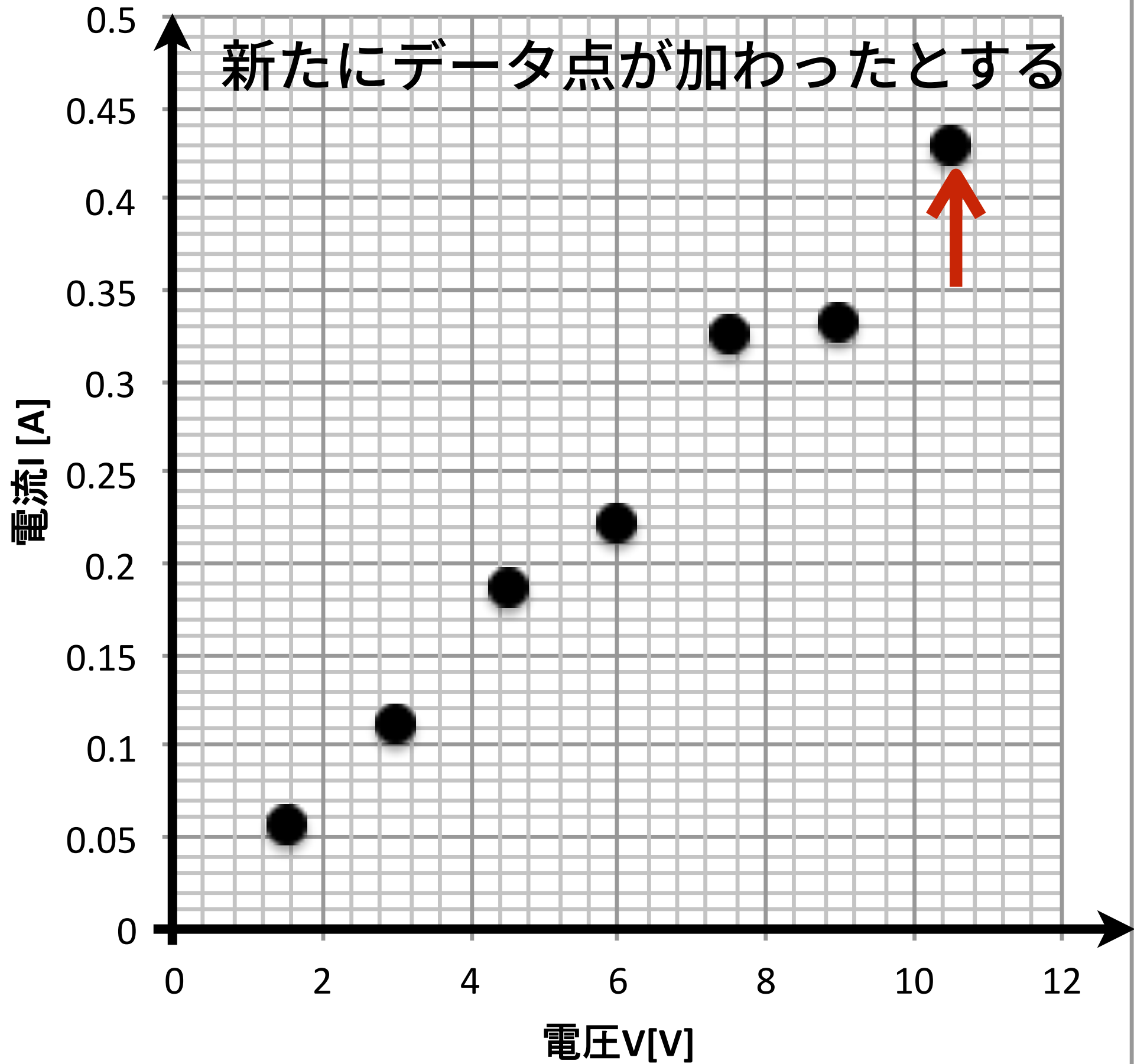
- ★ あるモデルを決めた時に，再現性の悪さを数値で表現できないか？
- ★ モデルの予測と，測定値のズレを見るのがよさそう
- ★ 全測定点に対して，ズレを足し合わせるのはいかがでしょうか？
  - ★ プラスとマイナスのずれが打ち消しあう可能性
- ★ しばしば二乗誤差が利用される。

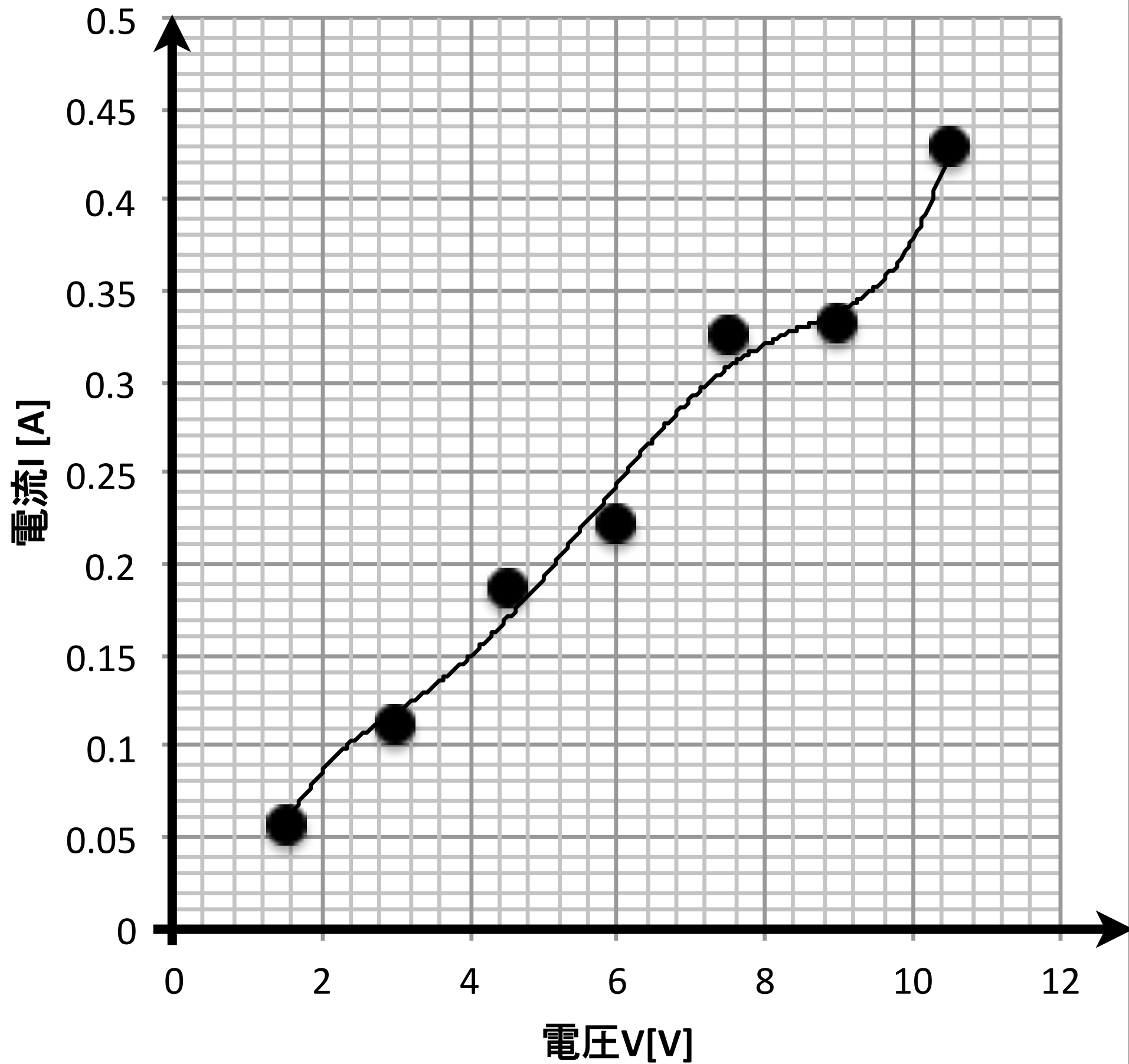
$I(V_i)$ はモデルの予測値。 $I_i$ は実測値

$$E = (I(V_1) - I_1)^2 + (I(V_2) - I_2)^2 + \dots$$

# 再現性の評価

- ★ 当然，二乗誤差の和が小さいモデルほど，データの再現性は良い。
- ★ モデルを複雑にすれば，一般に再現性は向上する
- ★ 新たな測定点が追加された場合，どの程度モデルを変更する必要があるか？
- ★ パラメータが多すぎるモデルの場合，根本から変更する必要がある出てくることも多い。





$$I = -0.230 + 0.352V - 0.148V^2 + 0.031V^3 - 0.00295V^4 + 0.000103V^5$$

# 比べてみよう

## 1点追加前

$$I = 0.667 - 0.928V + 0.492V^2 - 0.113V^3 + 0.0119V^4 - 0.00047V^5$$

## 1点追加後

$$I = -0.230 + 0.352V - 0.148V^2 + 0.031V^3 - 0.00295V^4 + 0.000103V^5$$

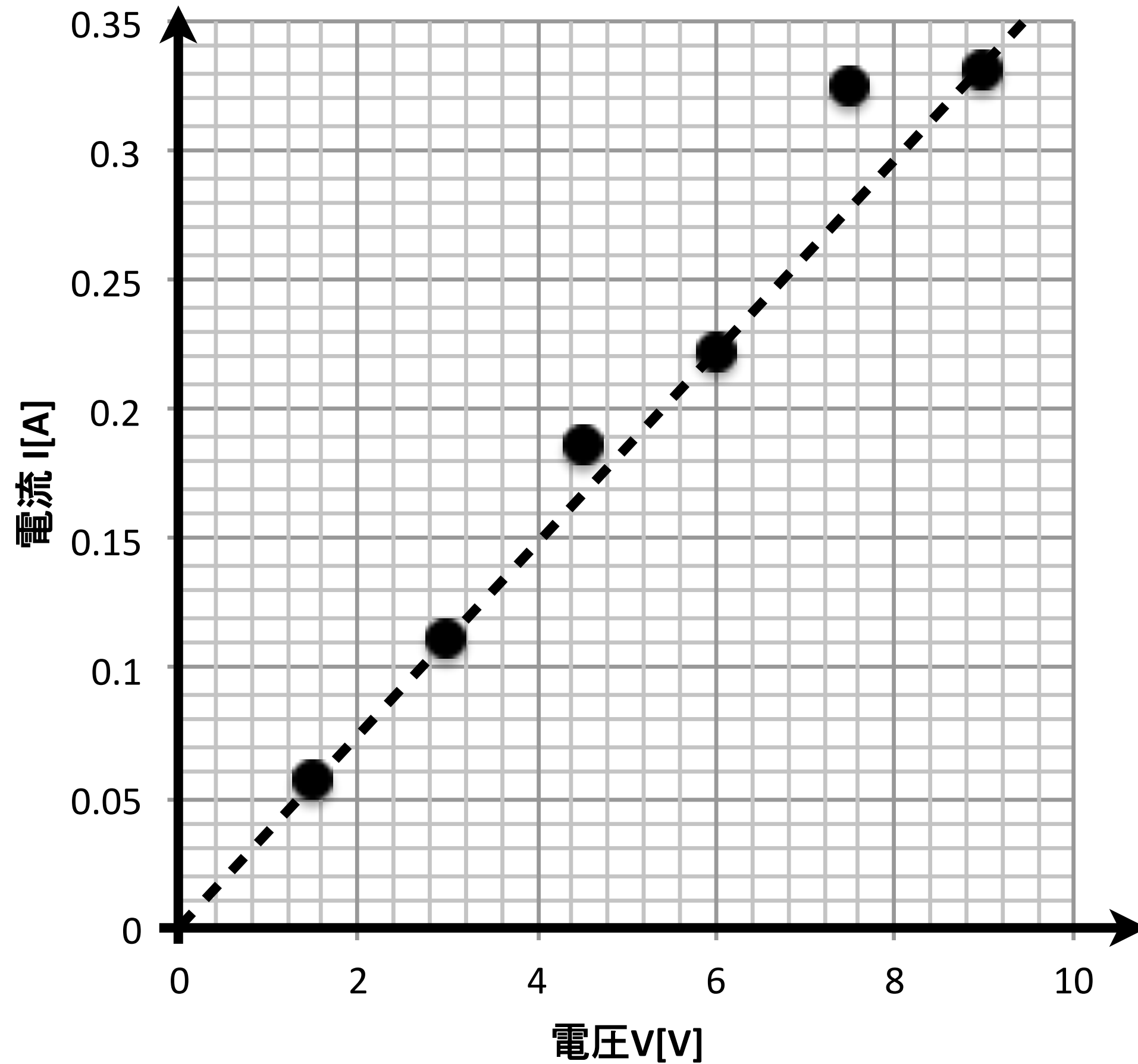
啞然とするくらい係数の値が変わった！！

5次関数のモデルはあまりいいモデルではなさそうだ。

# 演習問題

- ★ 前回作成したグラフをもとに，電圧 $V$ と電流 $I$ の関係を最もよく表すと思われる直線を「**点線で**」かきこめ
- ★ かきこんだ点線を  $I=aV$  という式で表すとする。傾き  $a$  を計算せよ。
- ★ この直線と測定データ間のズレを，二乗誤差を用いて評価せよ。



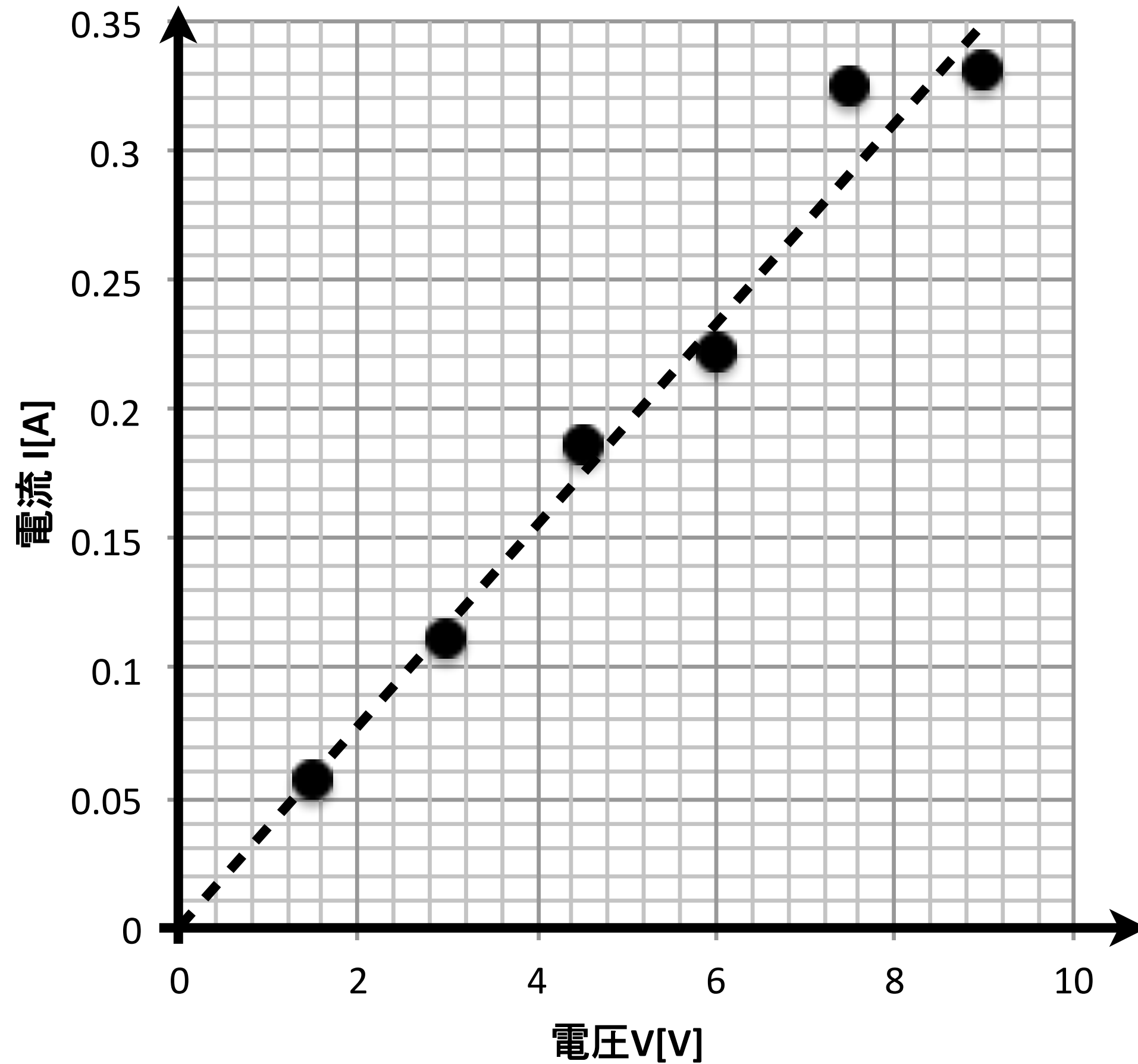


# 二乗誤差の計算例

電圧	1.5	3	4.5	6	7.5	9	
電流	0.0564	0.112	0.186	0.222	0.325	0.332	
予測値	0.0555	0.111	0.1665	0.222	0.2775	0.333	
差	-0.0009	-0.001	-0.0195	0	-0.0475	0.001	
二乗差	8.1E-07	1E-06	0.00038025	0	0.00225625	1E-06	0.00263931

$I = 0.037V$  の場合

$E=0.0026$



# 二乗誤差の計算例

電圧	1.5	3	4.5	6	7.5	9	
電流	0.0564	0.112	0.186	0.222	0.325	0.332	
予測値	0.0585	0.117	0.1755	0.234	0.2925	0.351	
差	0.0021	0.005	-0.0105	0.012	-0.0325	0.019	
二乗差	4.41E-06	2.5E-05	0.00011025	0.000144	0.00105625	0.000361	0.00170

$I = 0.039V$  の場合

$E=0.0017$

こっちの方が再現性は上