

自然科学の歩き方

第2回目

前回の学んだこと

- ★ データはうまく扱う必要がある
- ★ データをわかりやすく可視化するにはグラフにするのがよい
- ★ よくある質問
「点と点は線で結ぶんですか？」

ということで、今日は線の話をしてします

グラフ化の目的

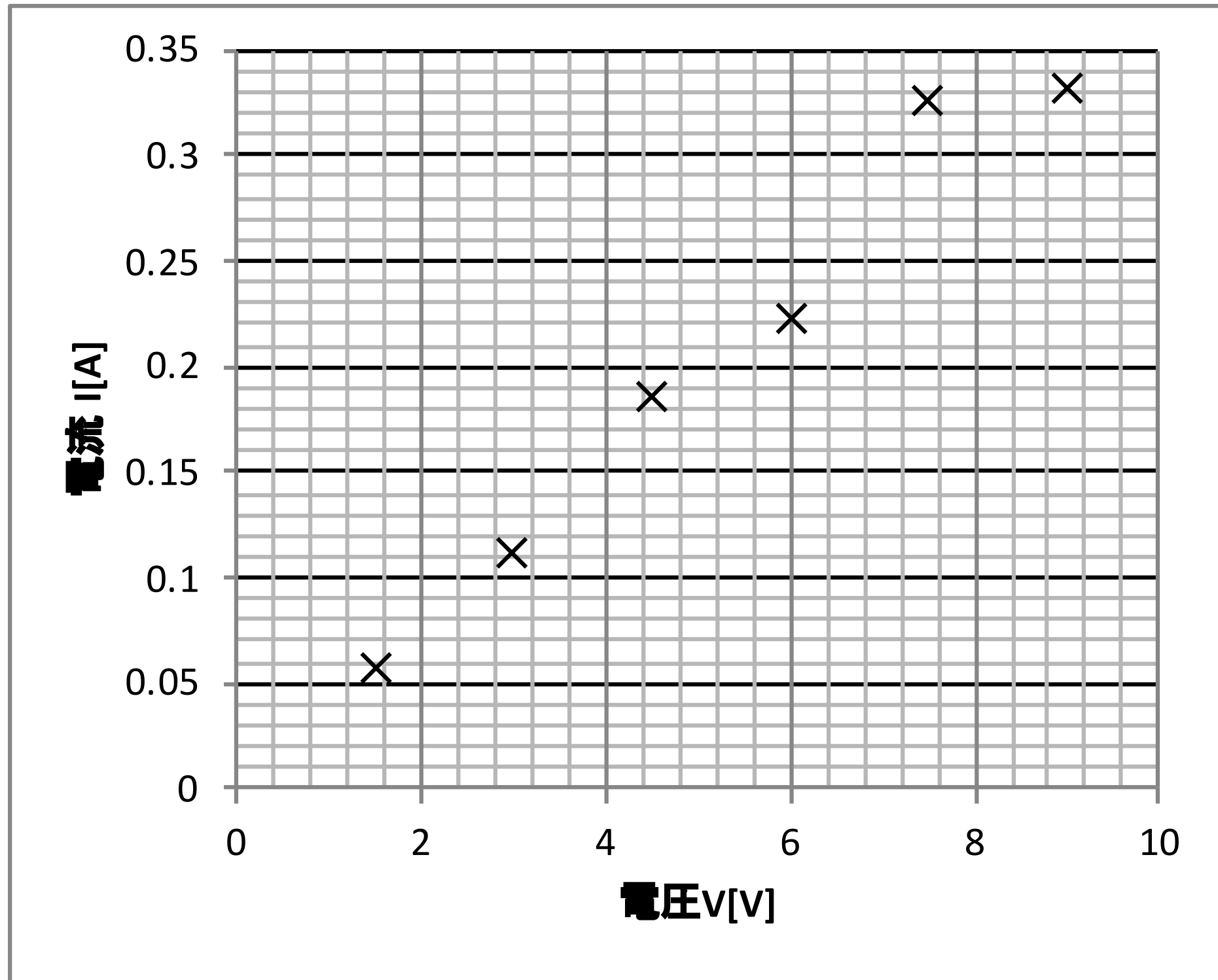
★ データの可視化

★ 何のためにやるか？

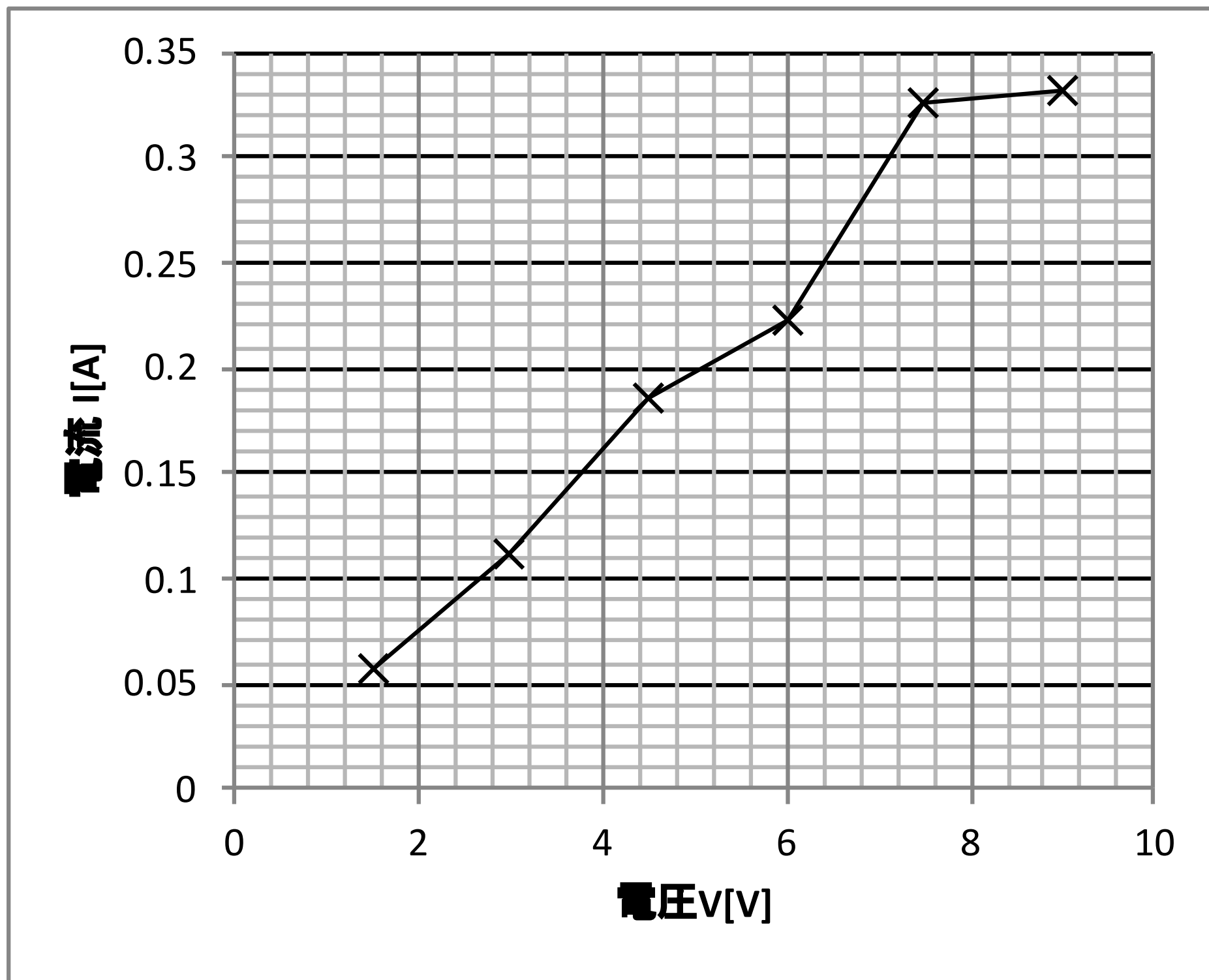
★ 見る人にわかりやすく示すため

★ 見た人がどう思うかを考える必要がある

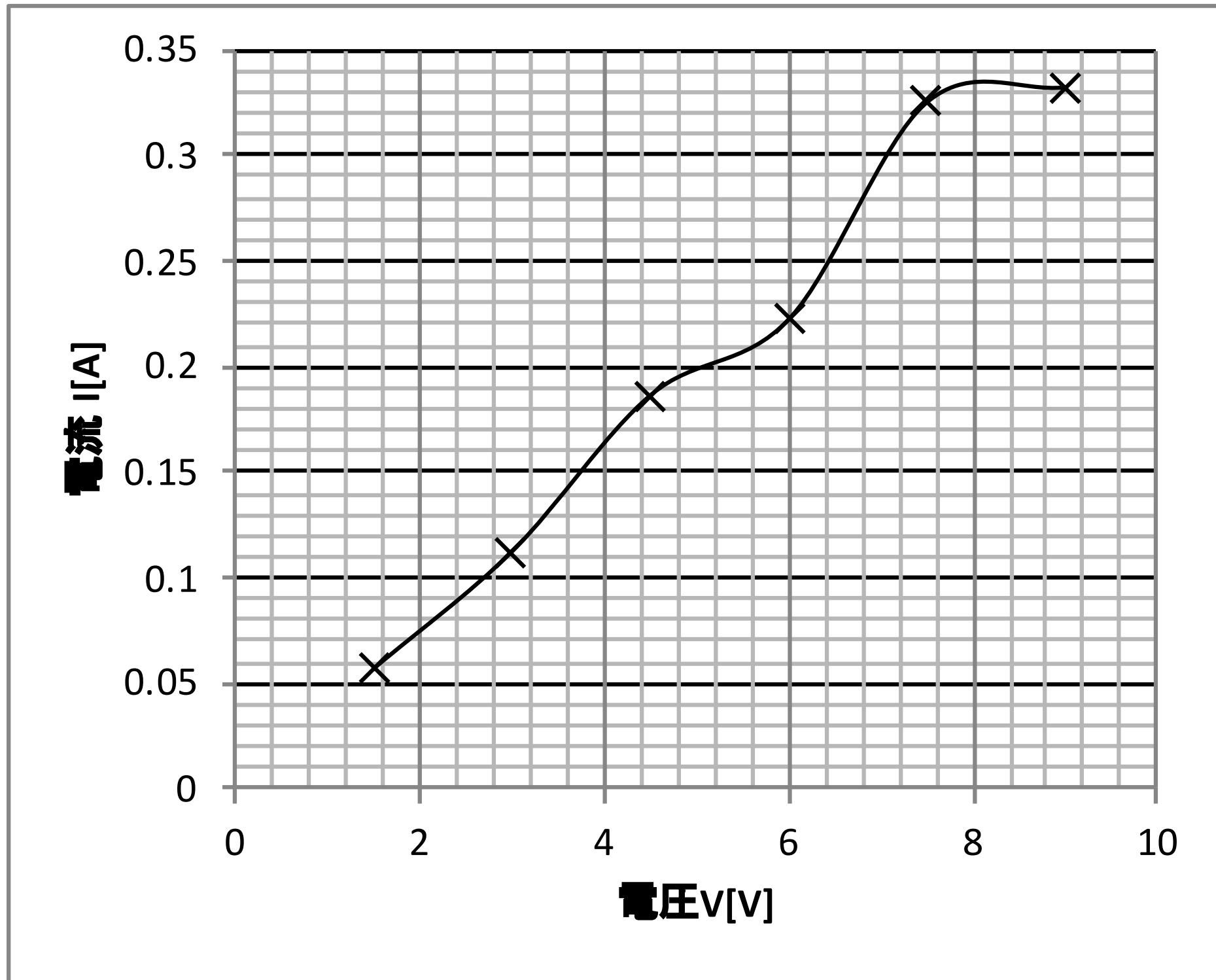
前回作ったグラフ



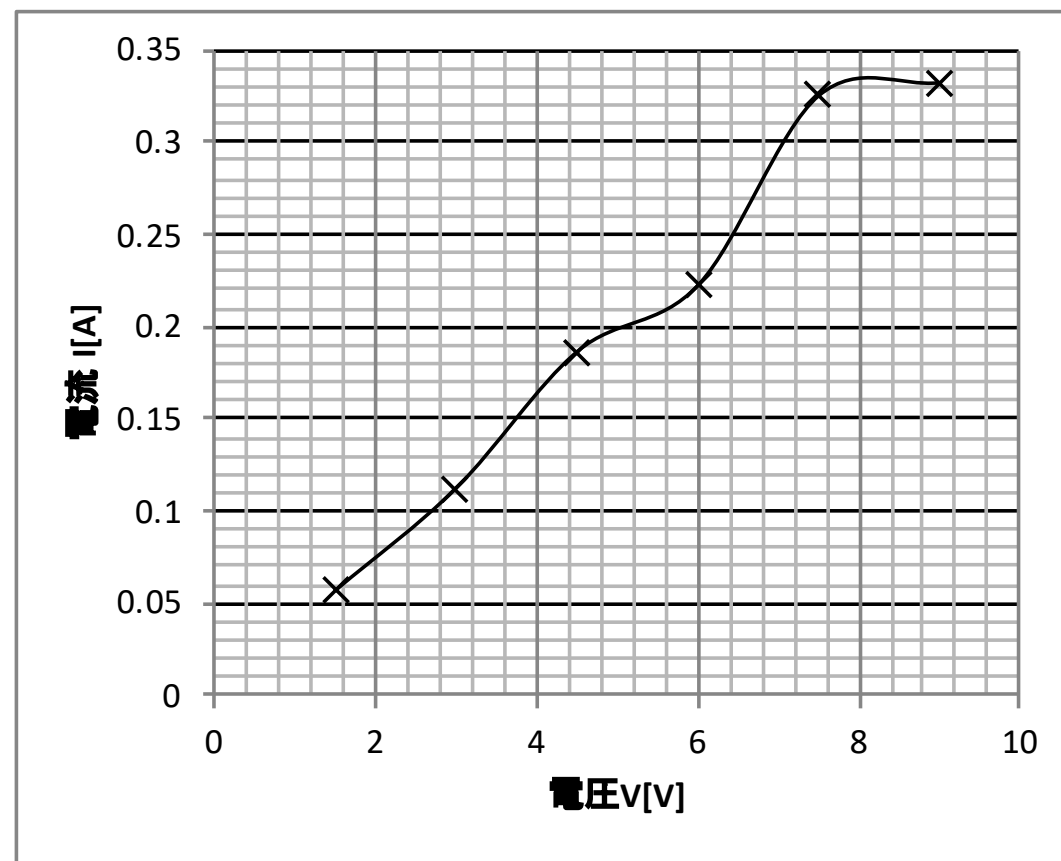
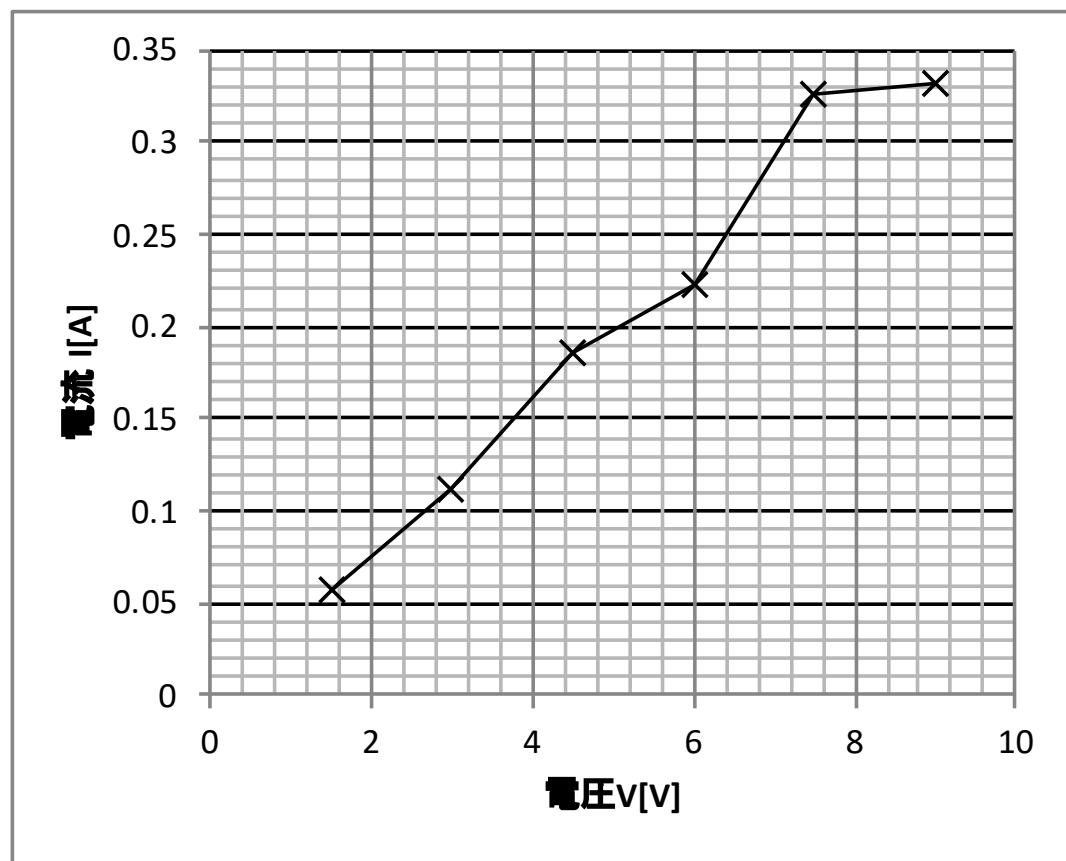
線でつないでみる



線でつないでみる 2



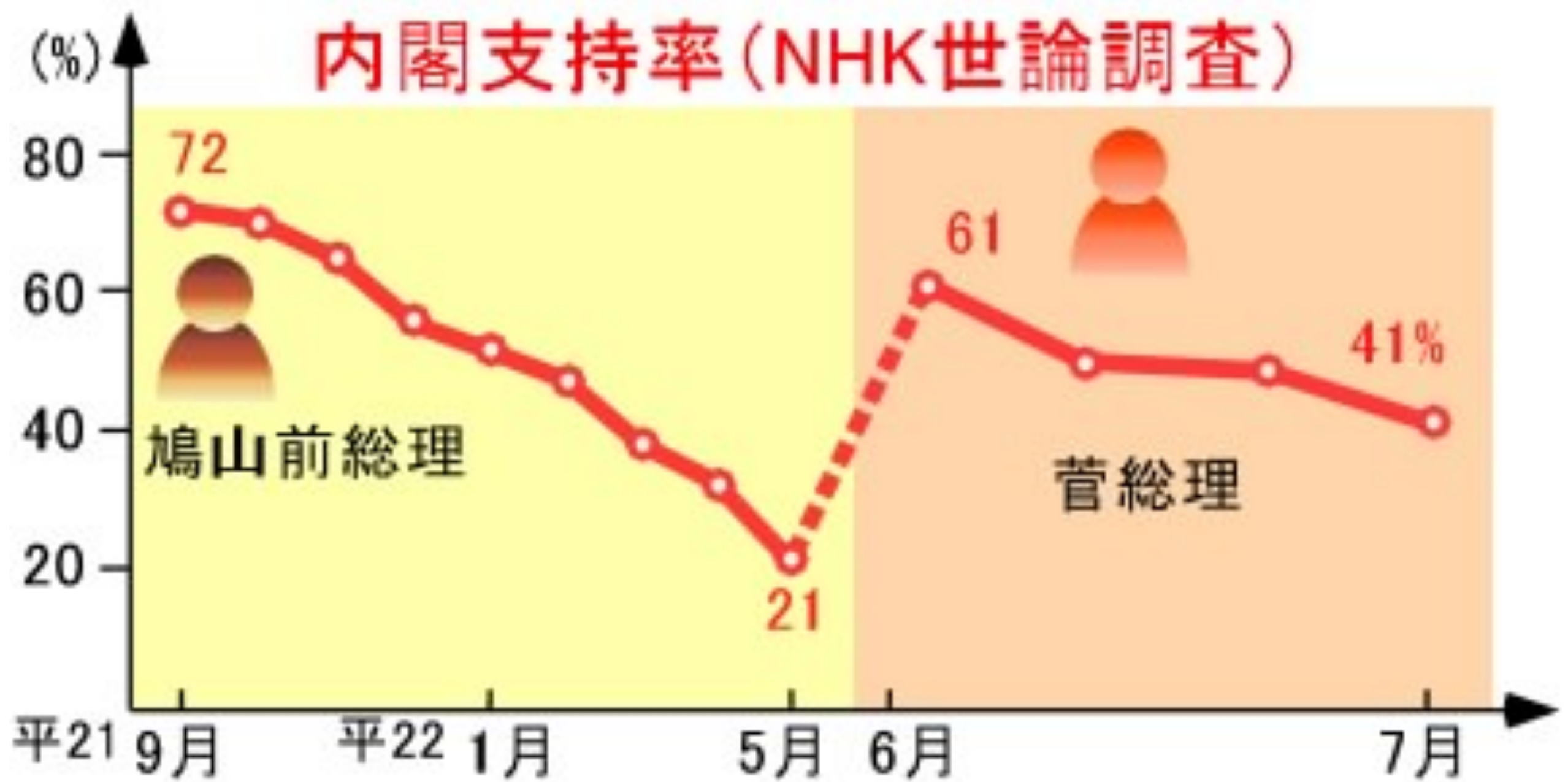
どう思いますか？



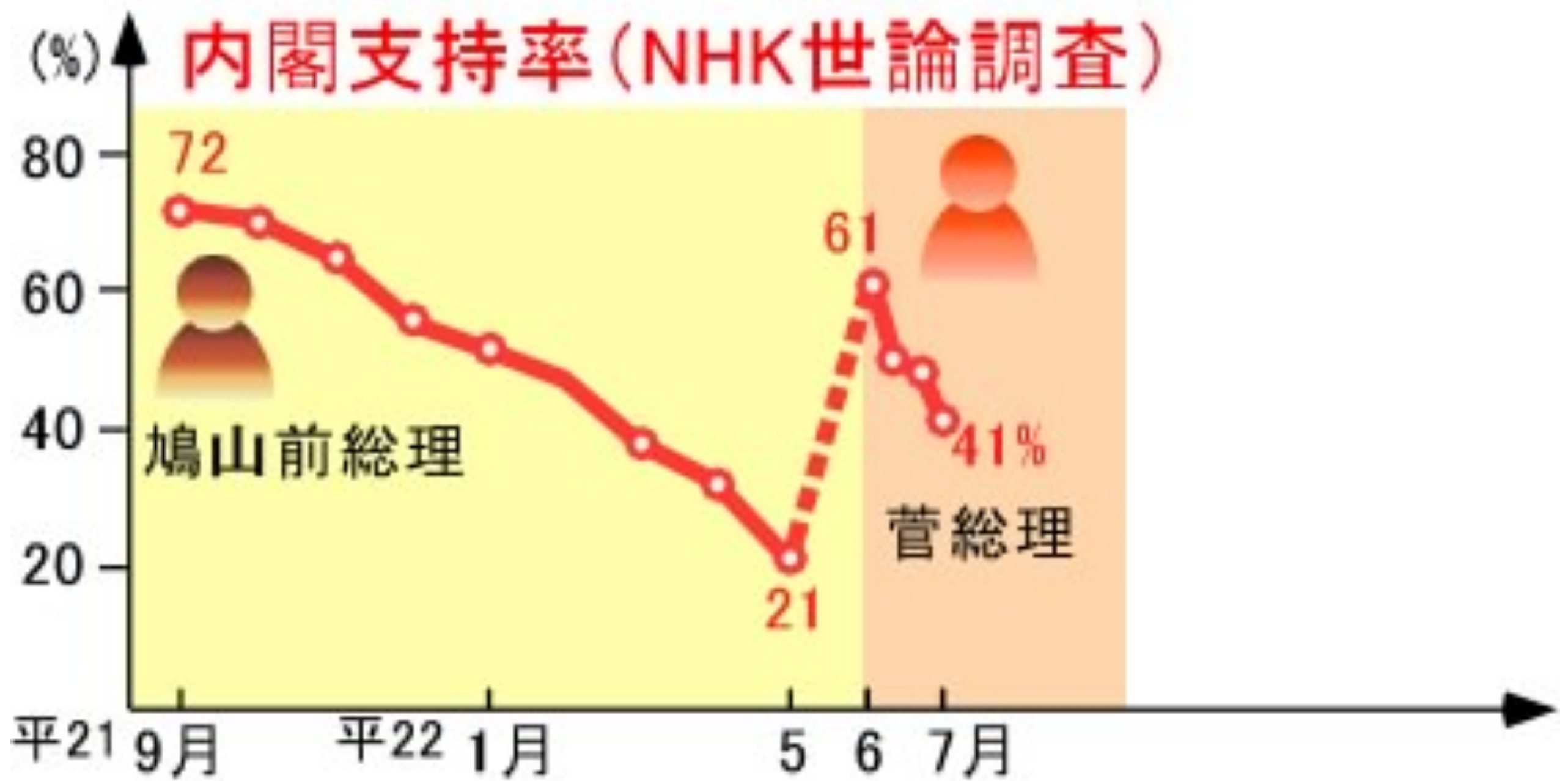
重要なこと

- ★ グラフは効果的にプレゼンするためのツールである
- ★ グラフに描きこむ「点」「線」「文字」には意味がある
- ★ 線は明確な意図をもって引かねばならない

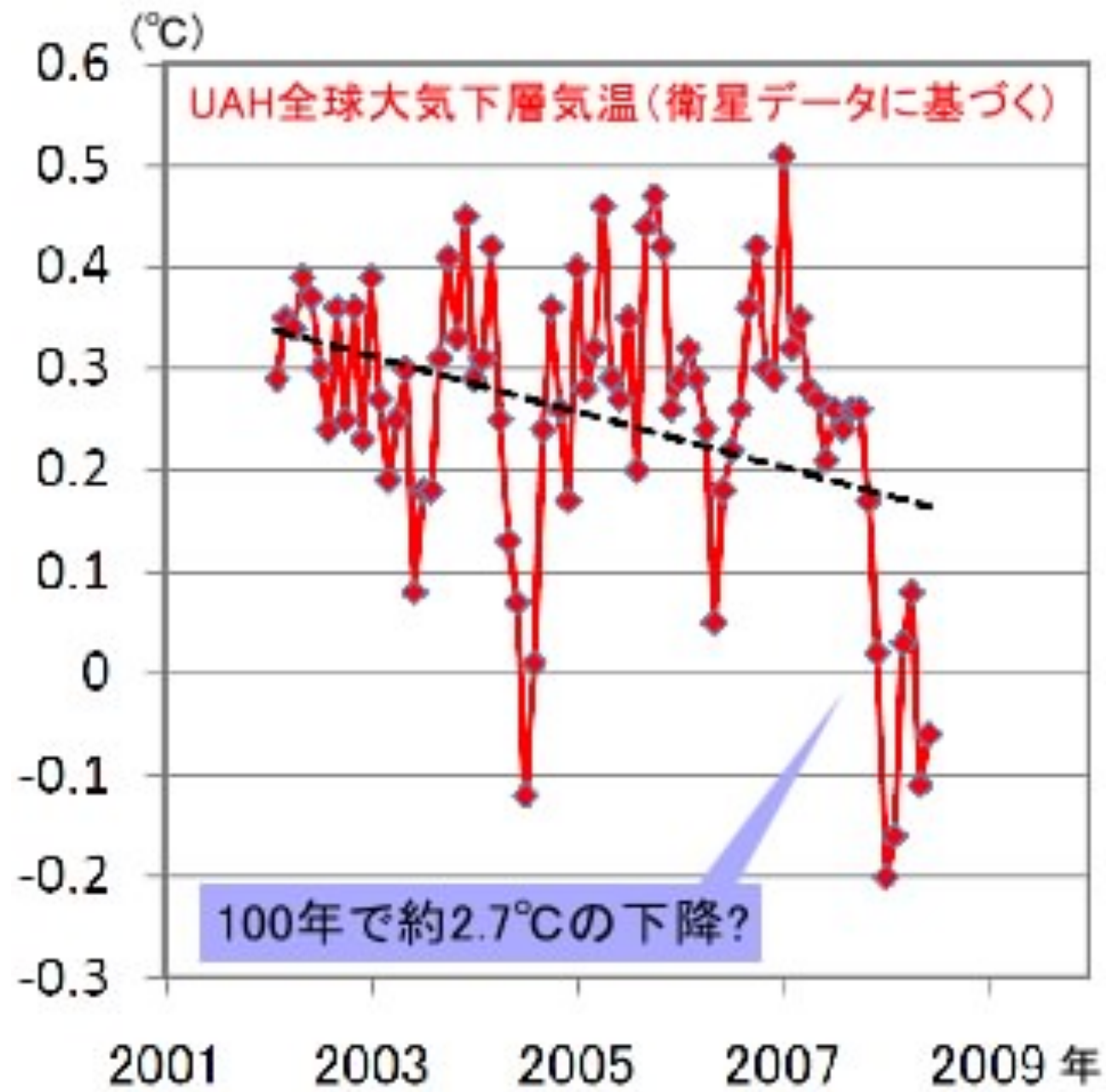
例



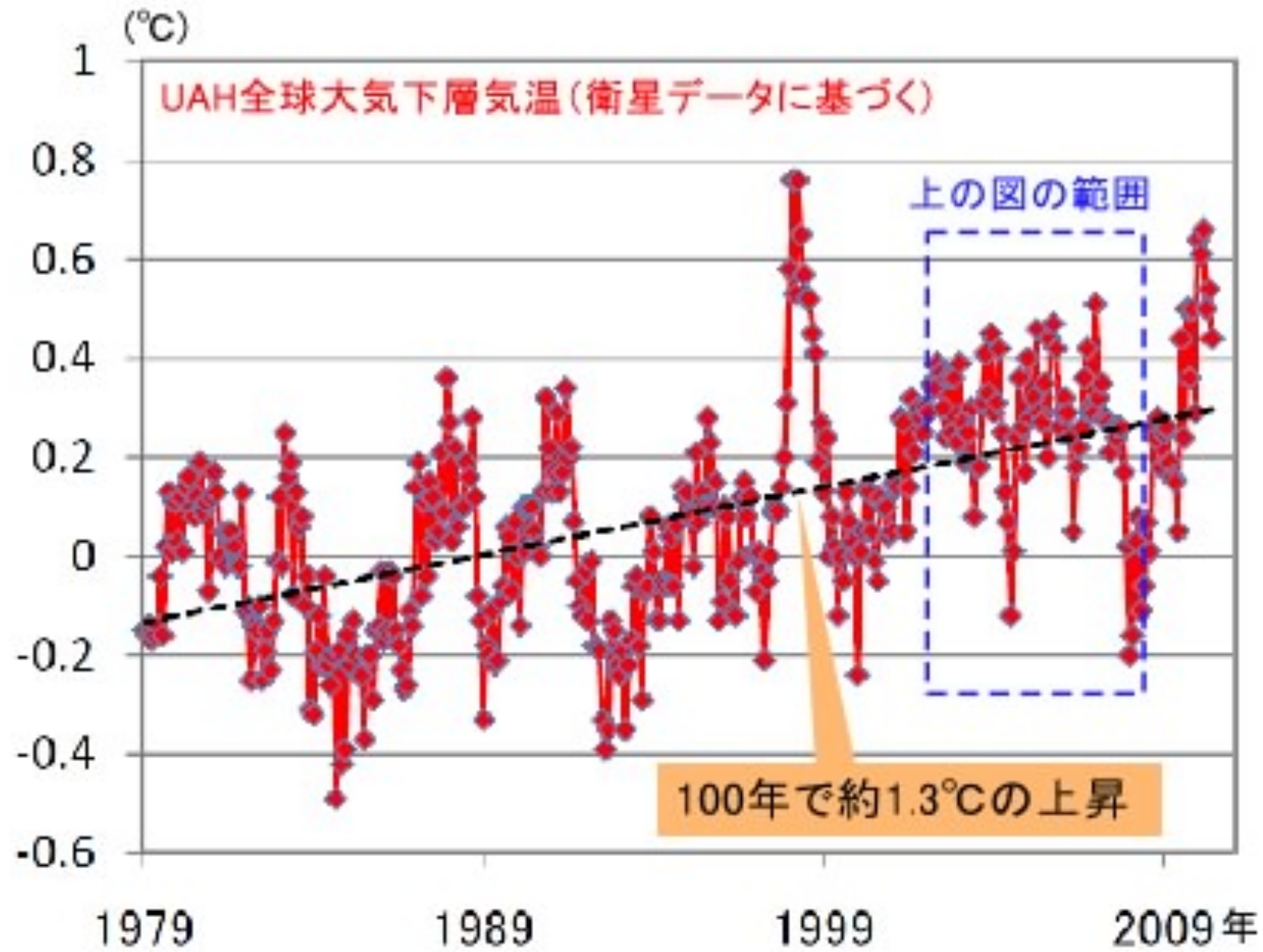
例



例 2



例



グラフを読む

- ★ データを視覚化したものがグラフである
- ★ グラフから意味のある情報を読み取る
 - ★ 数値データ→グラフ→言葉で表現
- ★ 視覚情報は思考に大きく影響する
 - ★ 何が読み取れるかをしっかり見極める
 - ★ 描くときには「何を伝えたいか」を表現

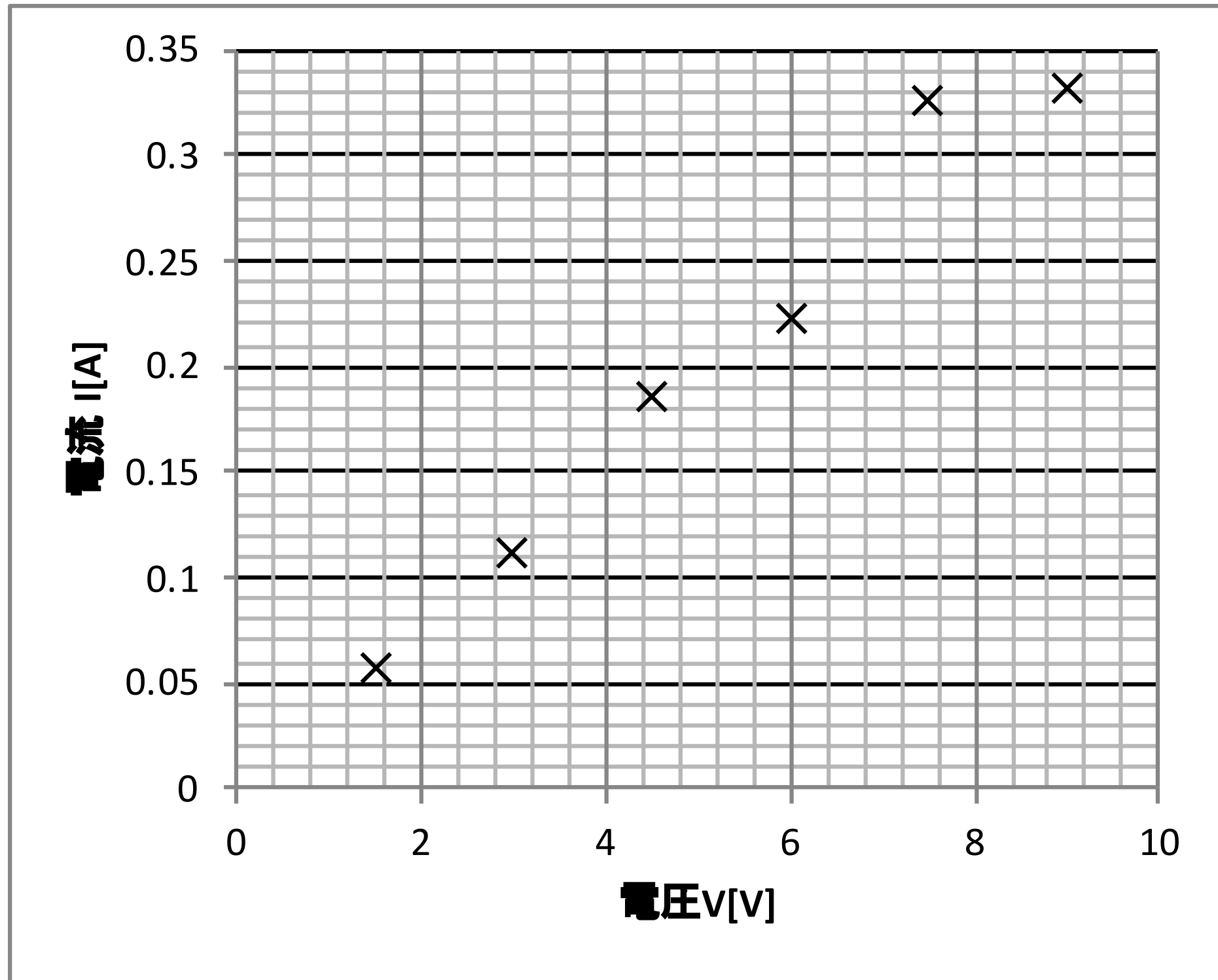
心構え

- ★ 自然に忠実に向き合うこと
- ★ 人間が知ることのできる情報は非常に限られているうう
- ★ データは自然の一部を切り取ってきたもの
 - ★ 自然そのものではない
 - ★ ある一面を示すのみ
 - ★ 誤差が含まれている

自然主義派

- ★ 測定したデータ点だけが真実
- ★ グラフにはデータだけ
- ★ どんなグラフが良いグラフか？

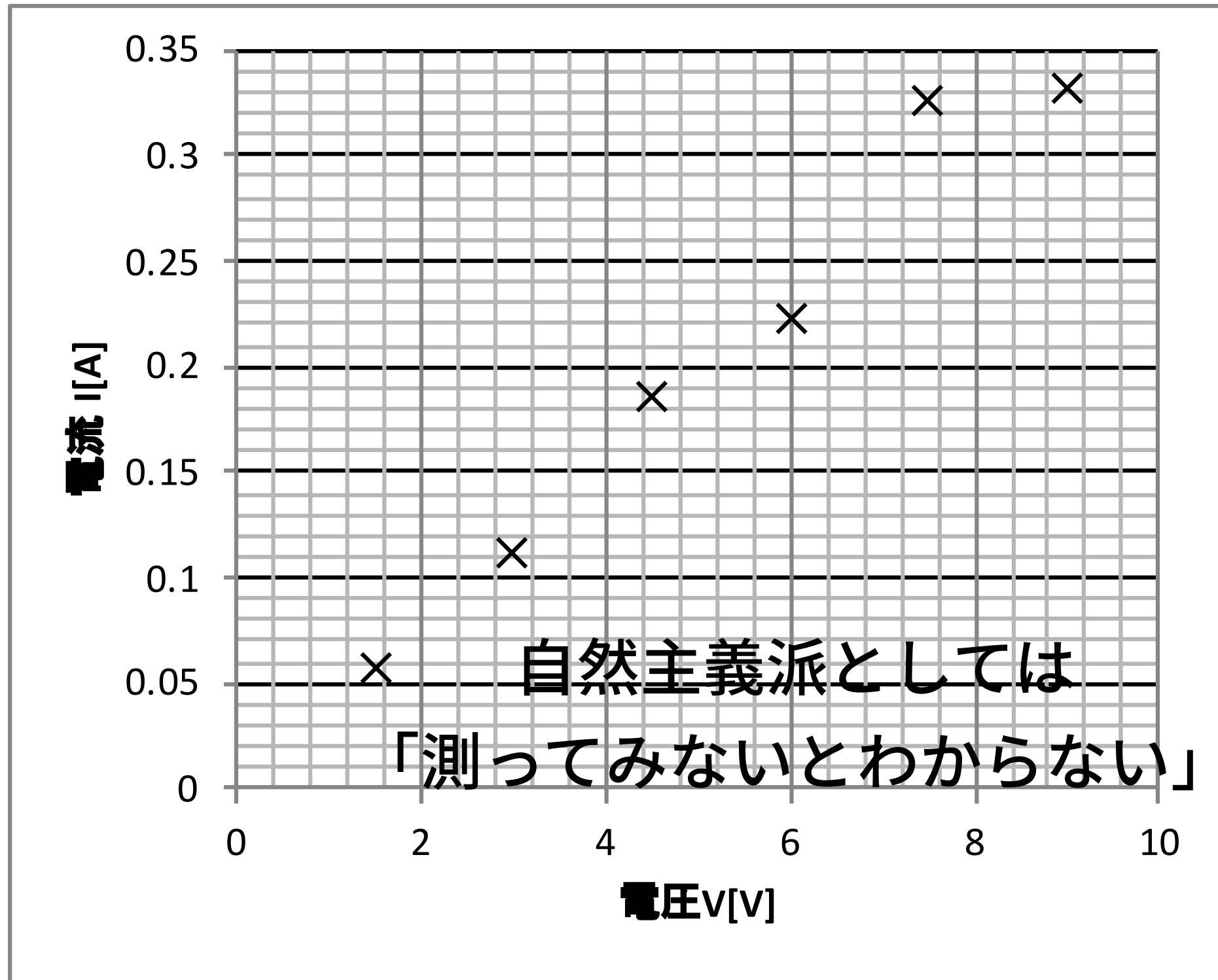
自然主義派のグラフ



自然主義派グラフの意味

- ★ データ点以外は信じない
 - ★ 測定していない点では何が起こるかわからない
 - ★ 何が起こるか知るには実験あるのみ
- ★ やや思考停止な感じ
- ★ 新しい現象の「予言」が重要
 - ★ 「わからない」では非常に効率が悪い
 - ★ とういか、いつまでも先に進まない

5Vのときの電流は？



仮説と予言

- ★ 電圧が大きくなると電流も大きくなる
- ★ どれくらい大きくなりそうか？
 - ★ 測定データを100%信頼
 - ★ 測定データをそれなりに信頼
- ★ データをもとに推定する

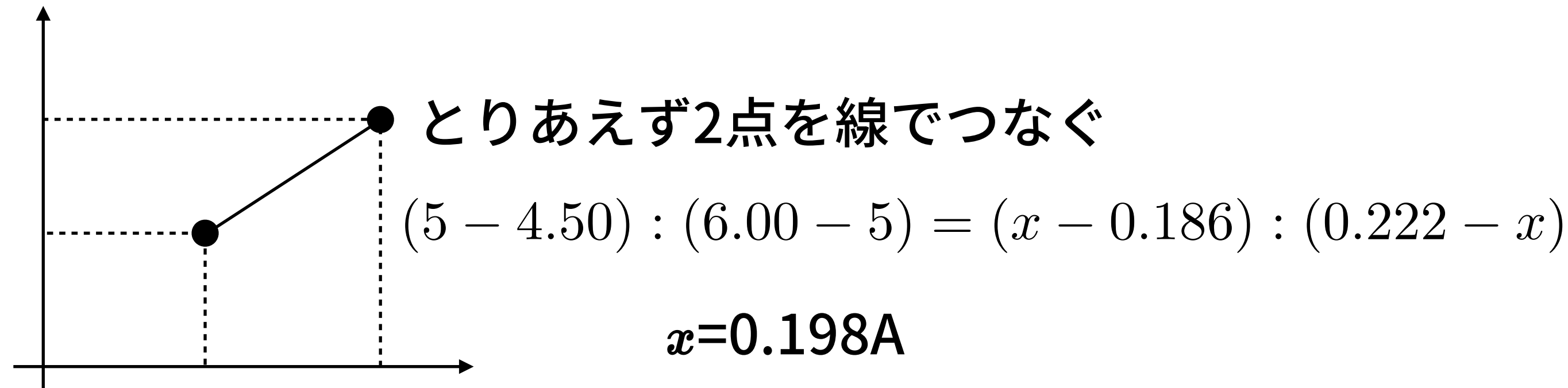
測定データを100%信じる

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

★ データ点をもとに，データの無いところに線をひくこと

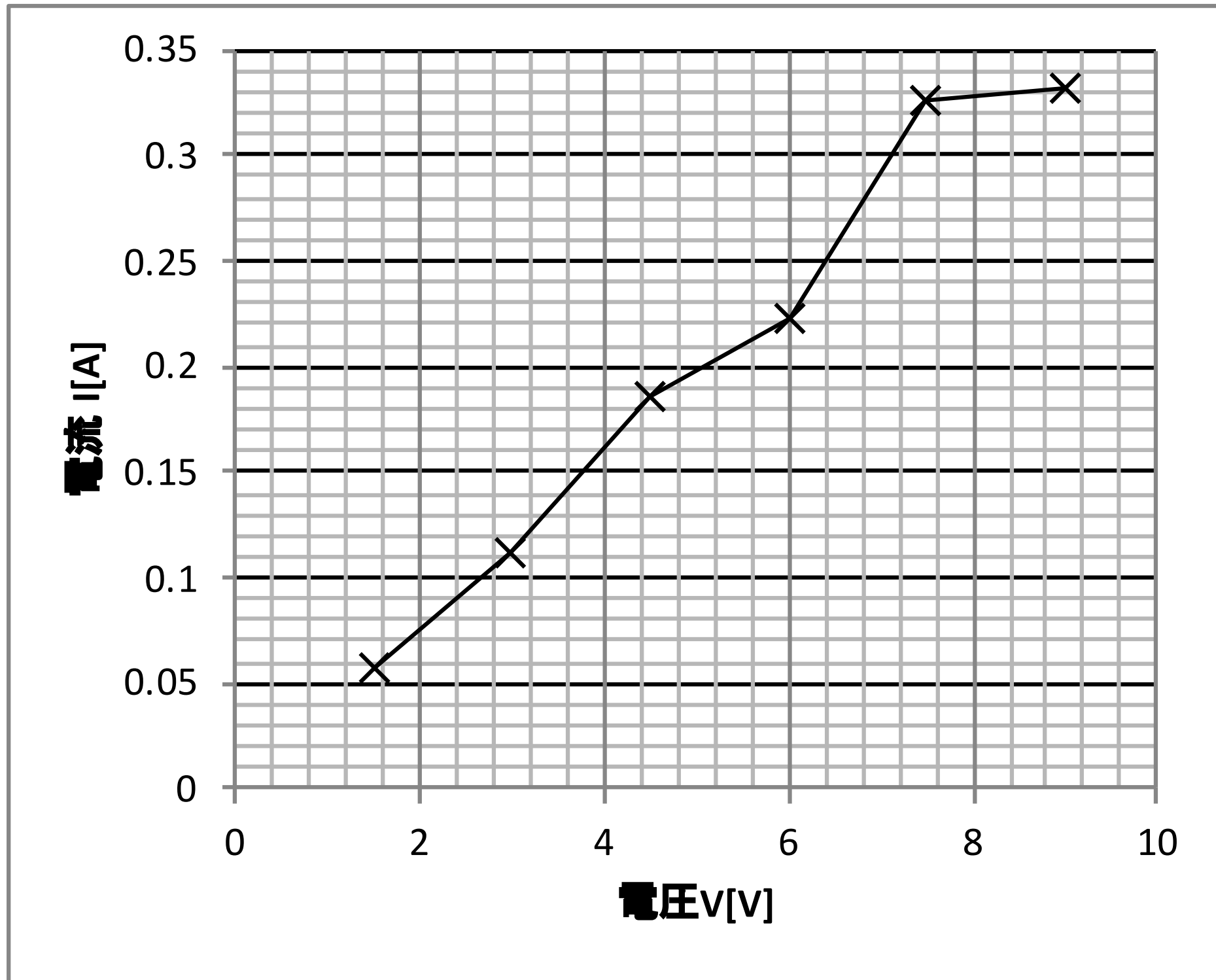
★ 電圧が4.50Vのとき，電流は0.186A

★ 電圧が6.00Vのとき，電流は0.222A



線型補完

非常に素朴



仮説・モデルの構築

- ★ データに基づいて仮説（モデル）をつくる
- ★ 仮説・モデルに基づいて予測を行う
- ★ 仮説・モデルが正しいかどうかは実験で確認

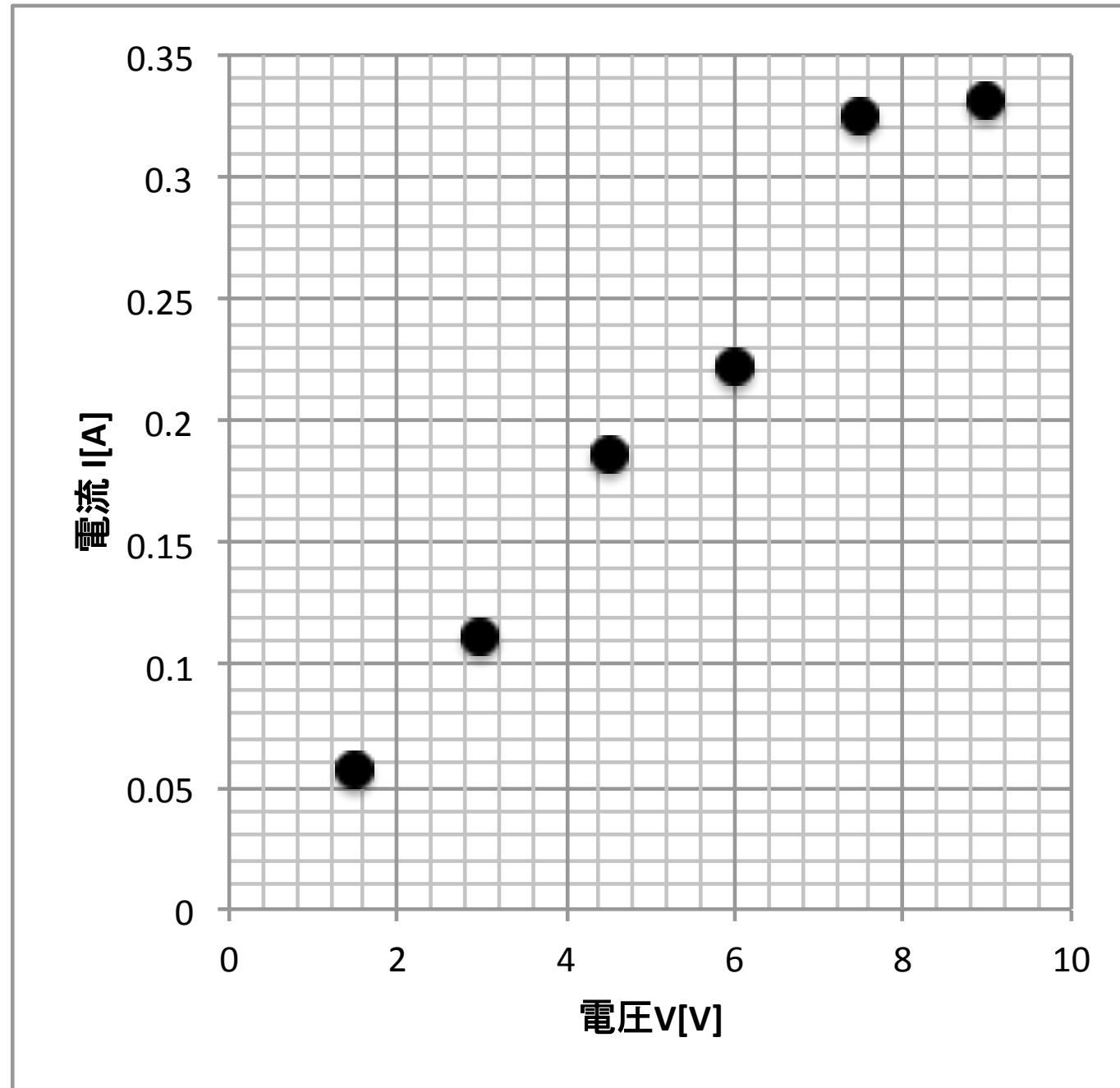
モデル

- ★ 「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

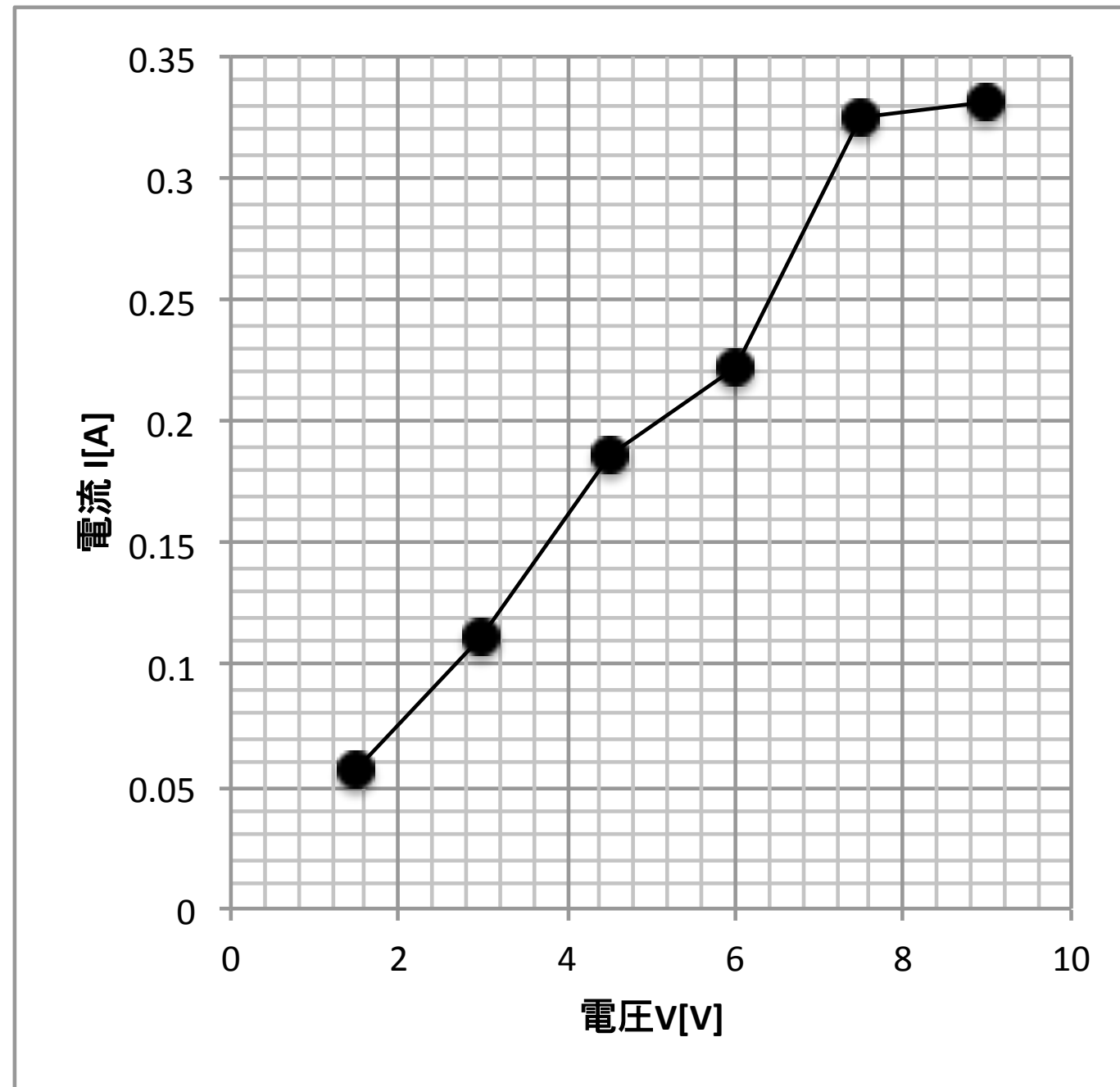
- ★ 全ての物理学（自然科学）の法則は「モデル」である
- ★ 言い換えれば、「自然科学は近似の学問」である
- ★ 「良いモデル」を作るには、枝葉末節を取っ払って、現象の本質を見極めることが重要になる
- ★ 質点の運動
- ★ 自由落下
- ★ ...

電流と電圧の例



データそのもの

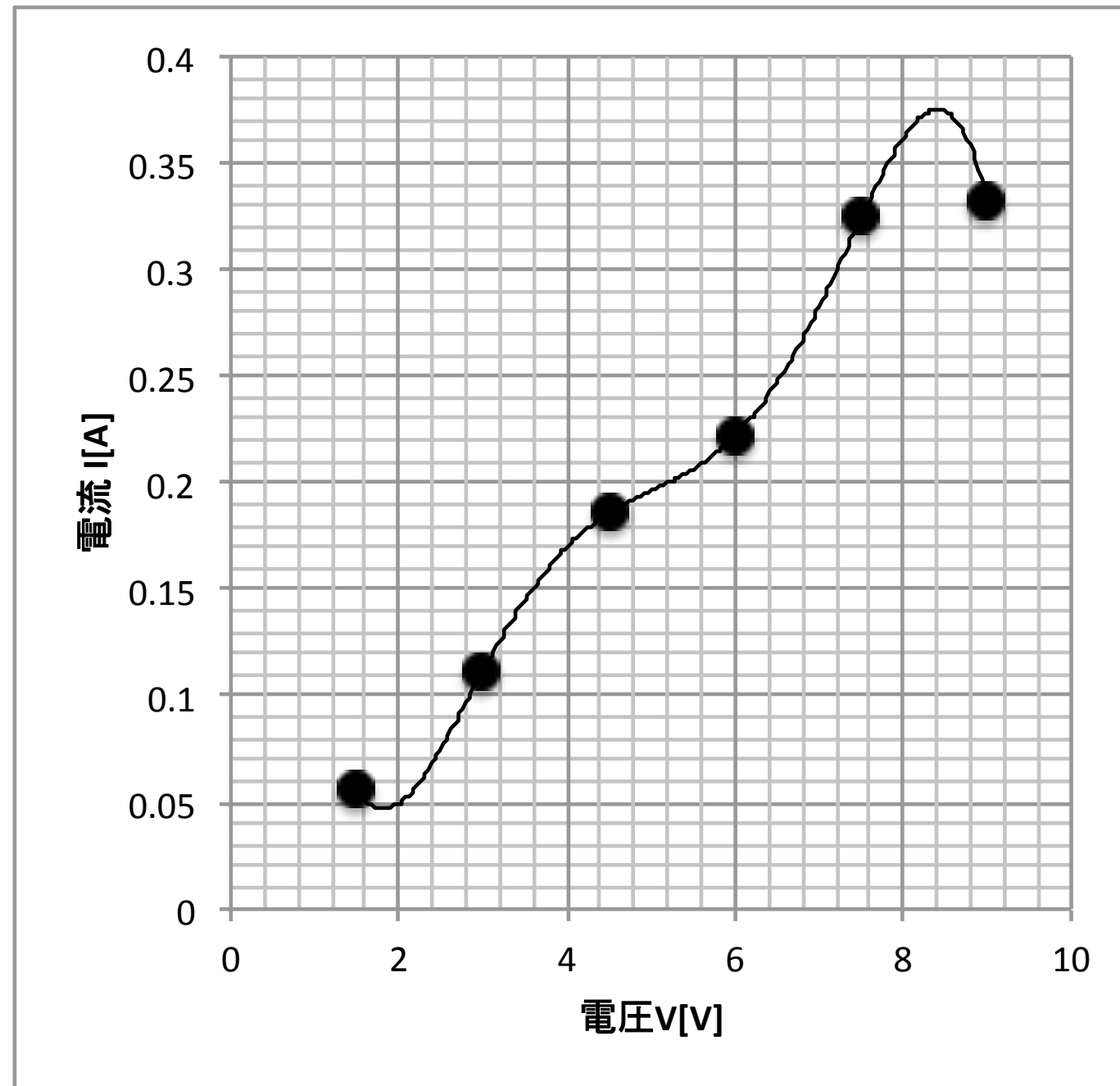
電流と電圧の例



線型補完もモデルの例

「電圧の区間ごとにふるまいが異なる」

電流と電圧の例



全部の領域を一つの式で表す

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

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

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

電流と電圧の：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？

$$I = aV$$

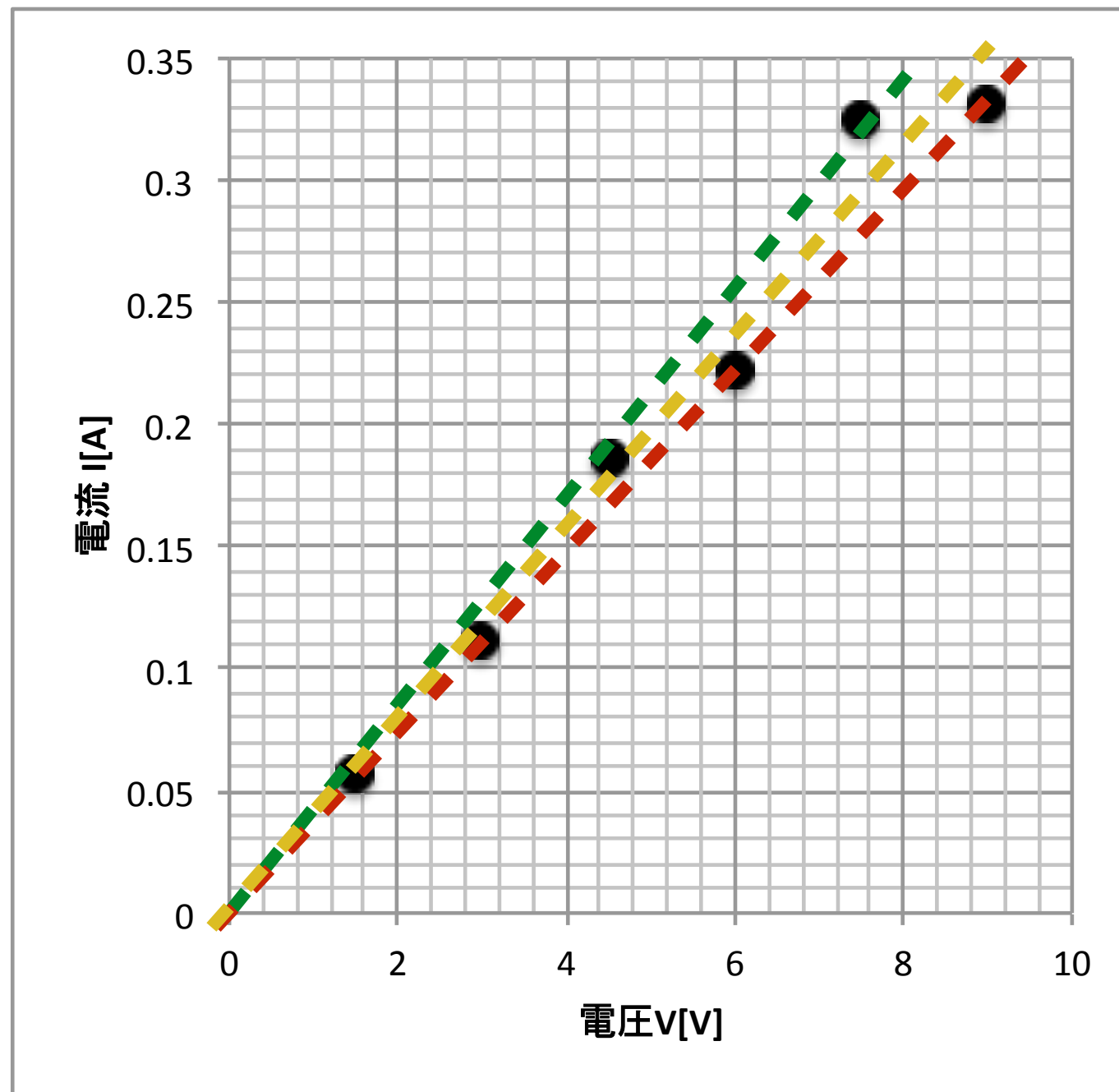
★ パラメータは1つだけ

★ 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

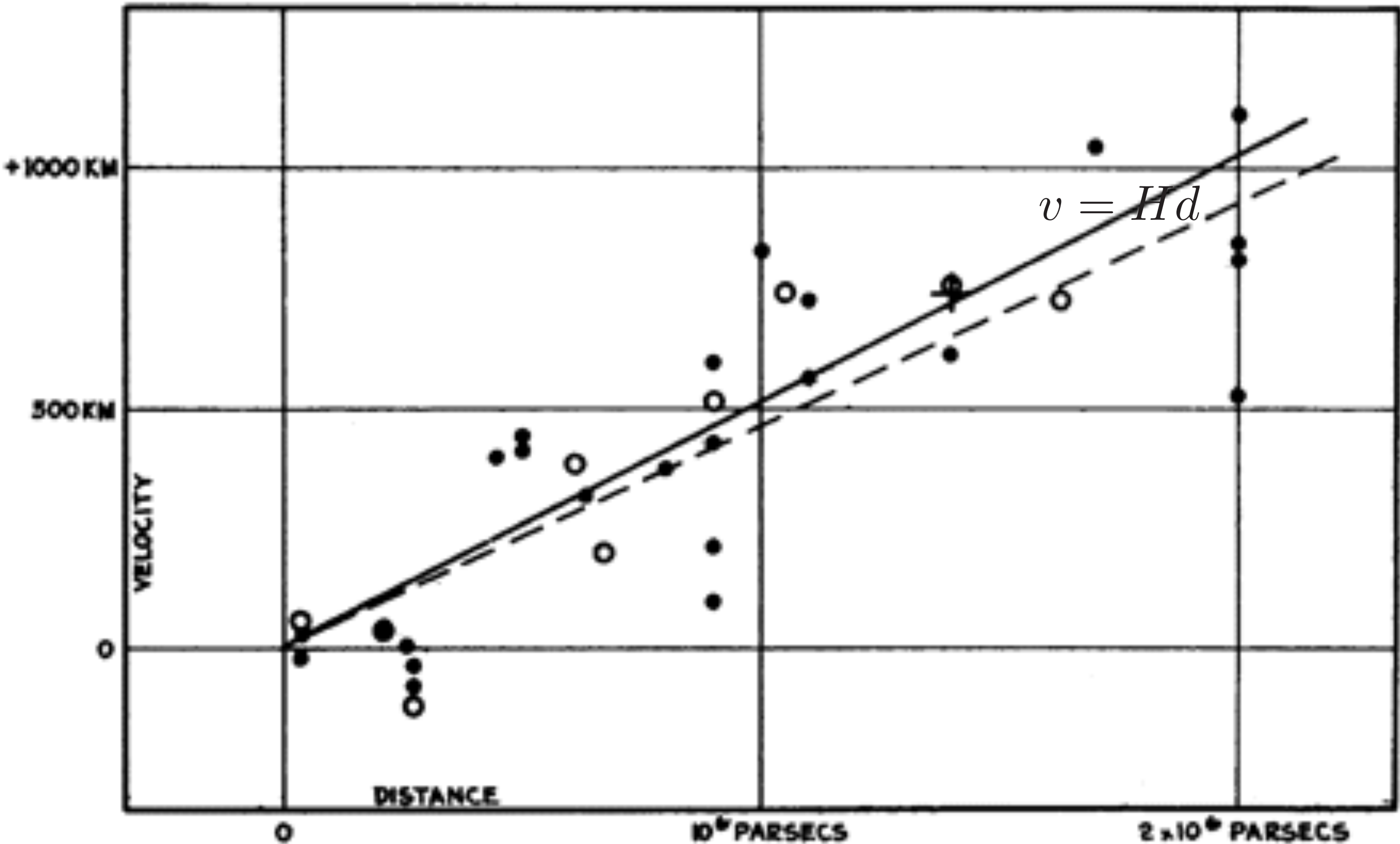
★ 傾きの値をどう設定するべきか？

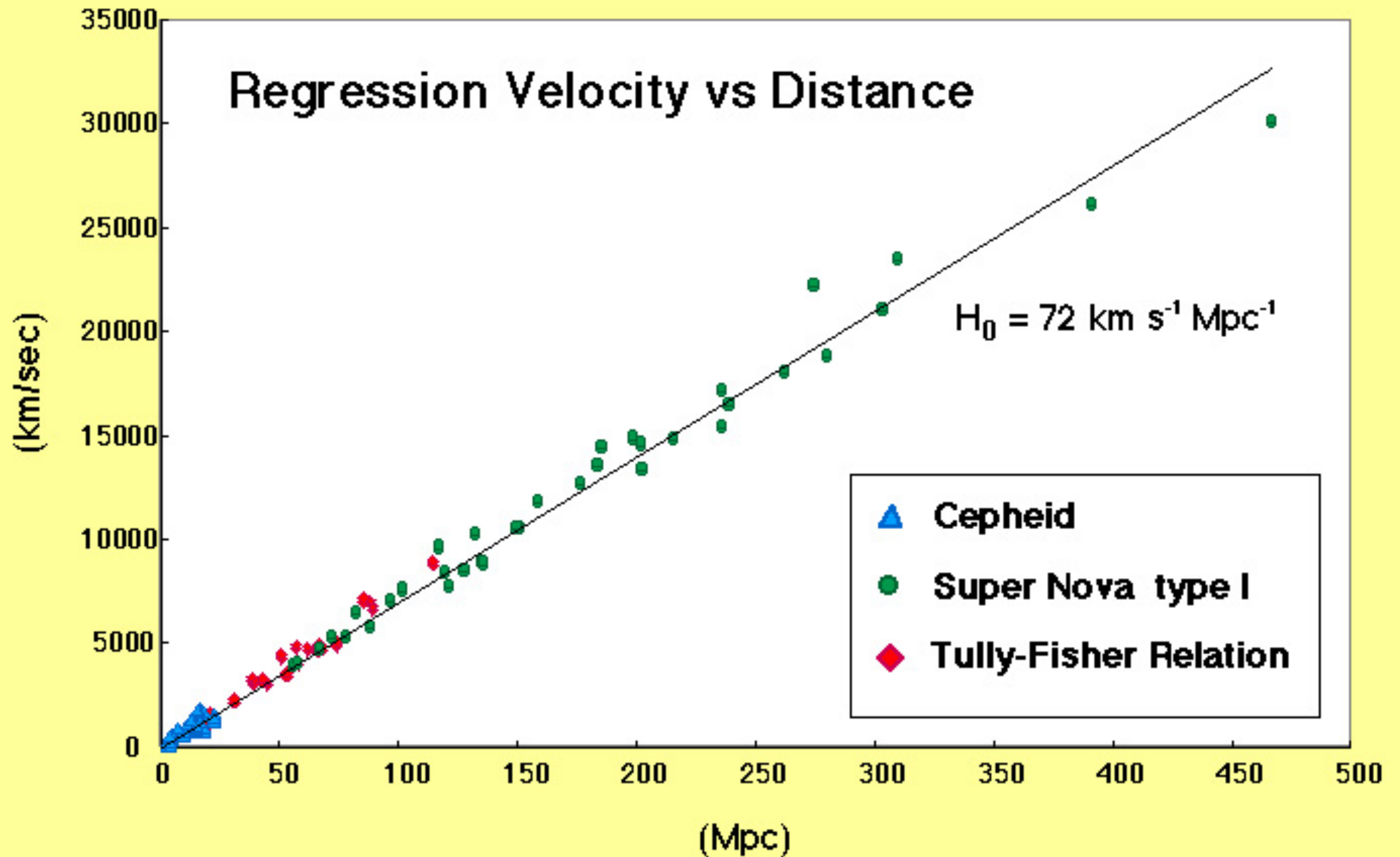
電流と電圧の例



シンプルに

E. Hubble, Proceedings of the National Academy of Sciences of the United States of America **15** (3): 168





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

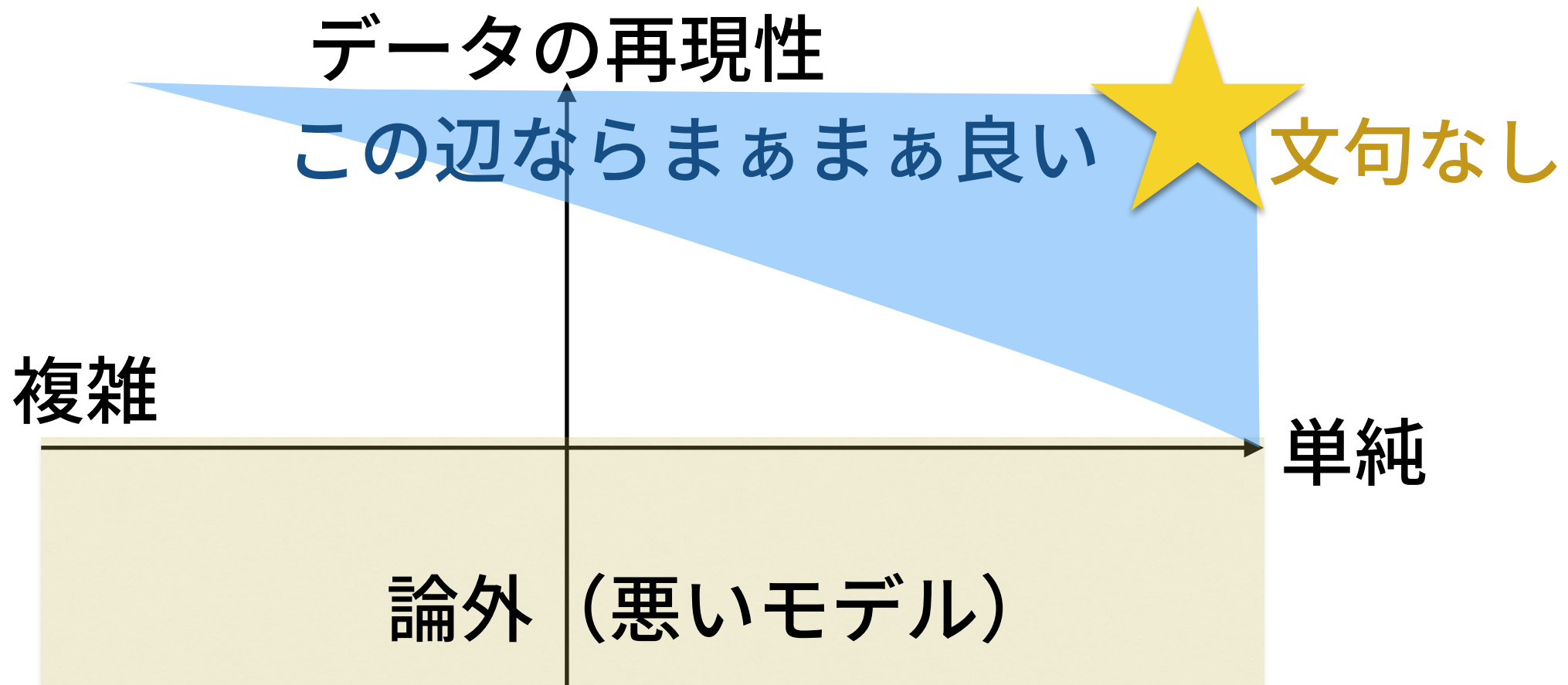
$$I = c_0 + c_1V + c_2V^2 + c_3V^3 + c_4V^4 + c_5V^5$$

★ 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

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？



データの再現性を評価する

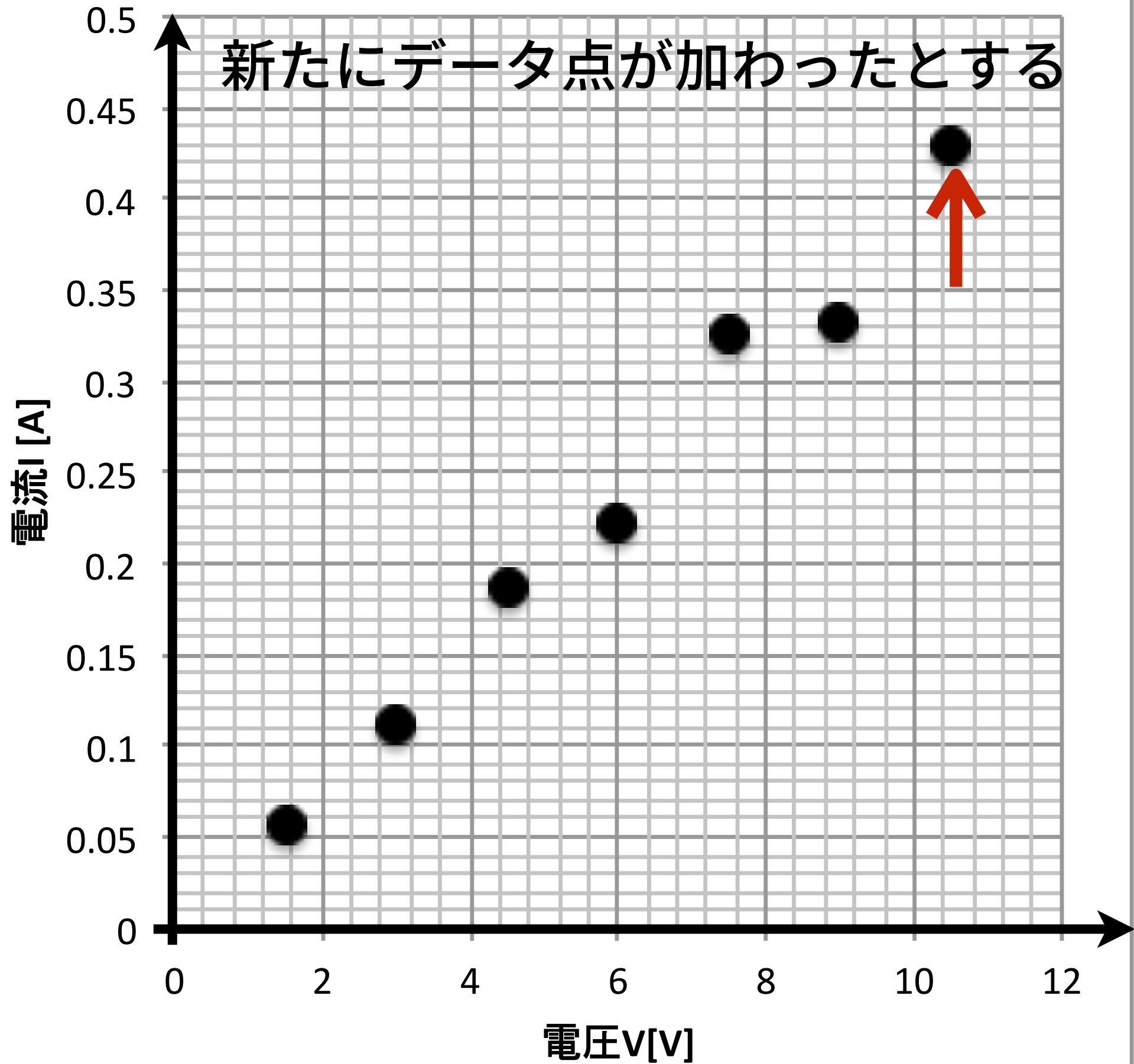
- ★ あるモデルを決めた時に，再現性の悪さを数値で表現できないか？
- ★ モデルの予測と，測定値のズレを見るのがよさそう
- ★ 全測定点に対して，ズレを足し合わせるのはどうか？
 - ★ プラスとマイナスのずれが打ち消しあう可能性
- ★ しばしば二乗誤差が利用される。

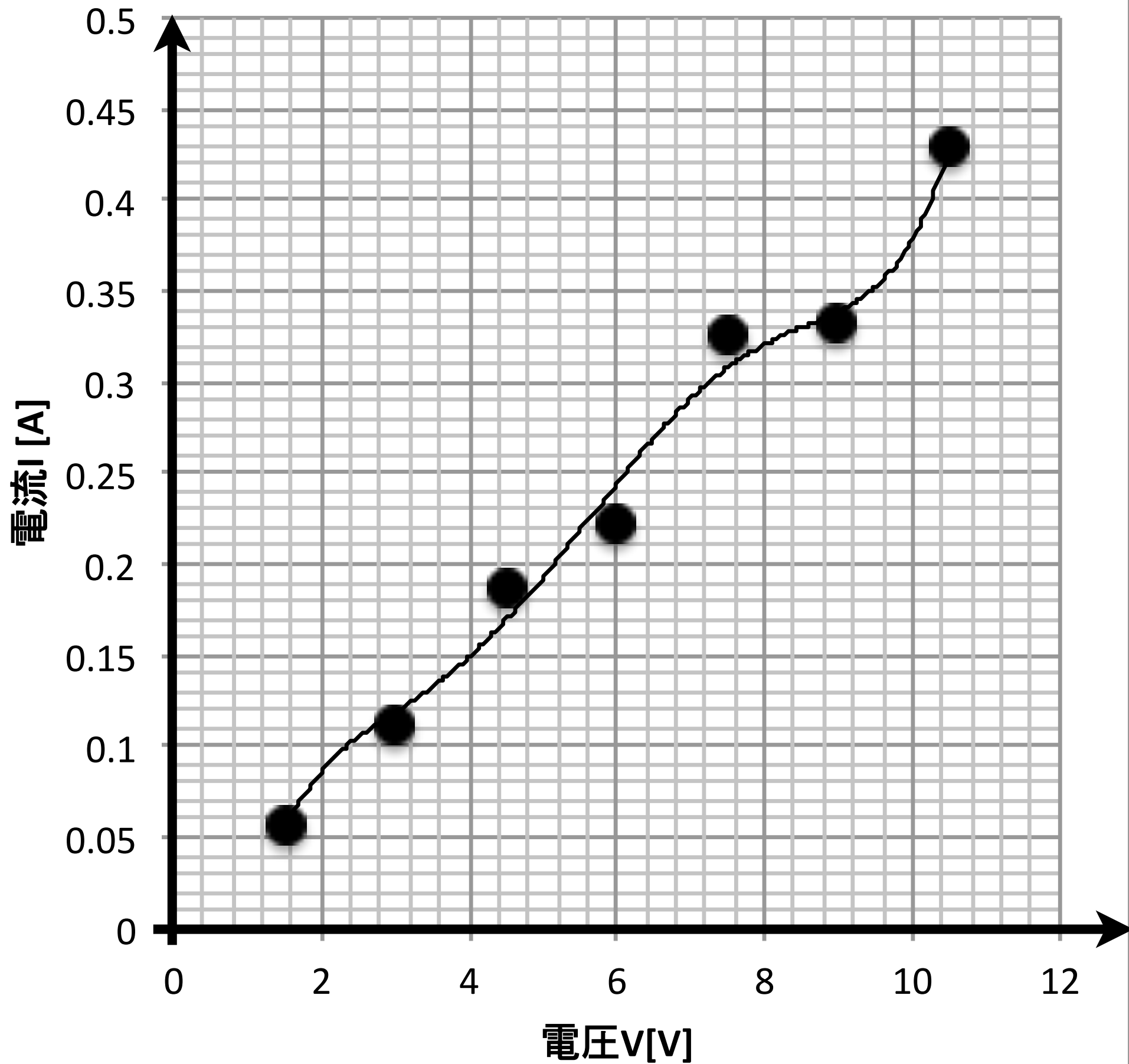
$I(V_i)$ はモデルの予測値。 I_i は実測値

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

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

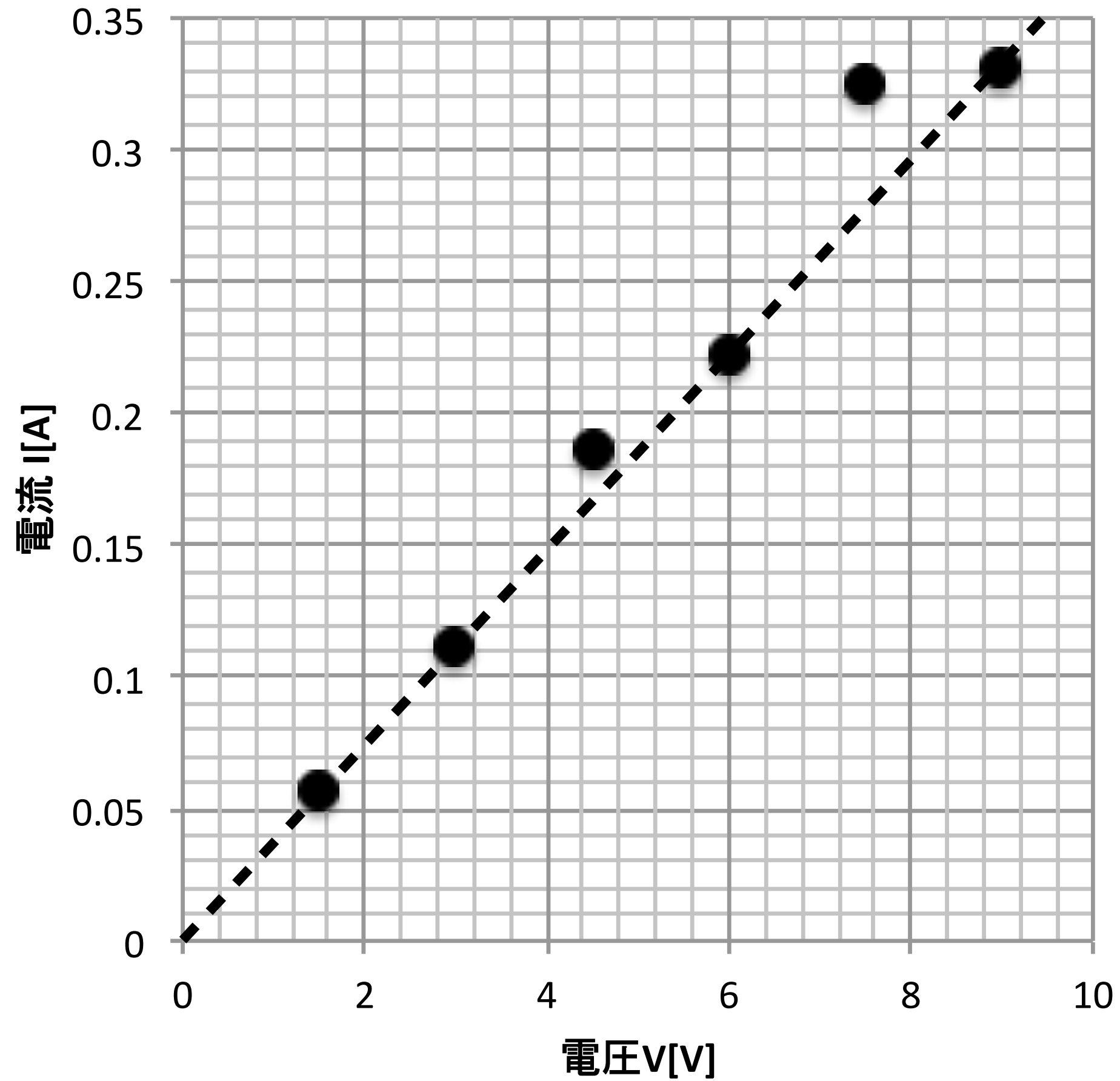
$$I = -0.230 + 0.352V - 0.148V^2 + 0.031V^3 - 0.00295V^4 + 0.000103V^5$$

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。

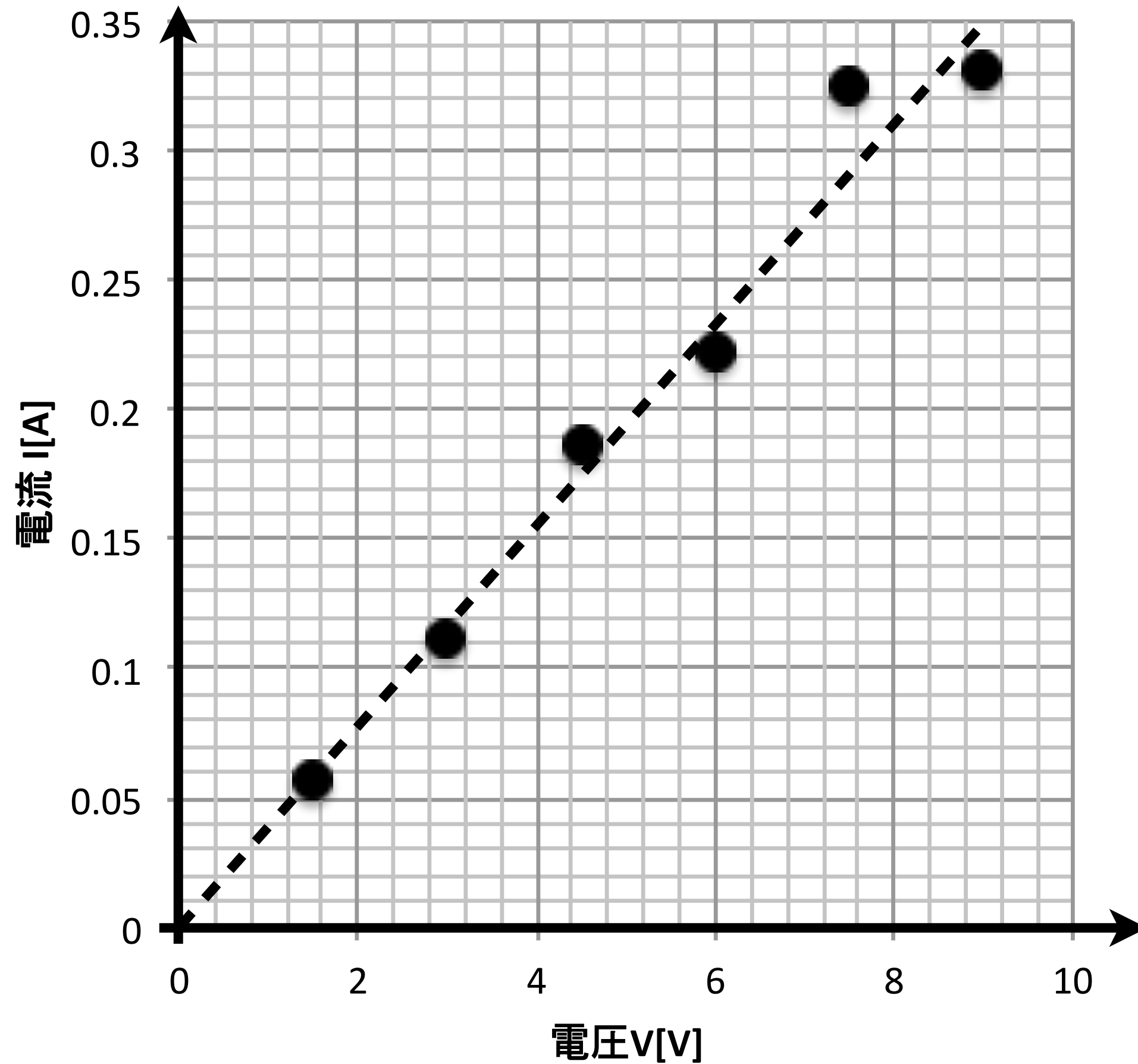


二乗誤差の計算例

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	0.000380	0	0.002256	1E-06	0.002639

$I = 0.037V$ 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	0.1112	0.186	0.222	0.325	0.332
予測値	0.0585	0.1117	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	0.000110	0.000144	0.001056	0.000361

$I = 0.039V$ の場合

$E=0.0017$

こっちの方が再現性は上