標本から母集団のパラメータ（母数）を推定する際、推定量が「不偏」であるとはどういうことだろうか？

不偏推定量の定義

推定量 $\hat{\theta}$ が母数 θ の**不偏推定量**であるとは、

$$\mathrm{E}(\hat{\theta}) = \theta$$

が成り立つことをいう。

つまり、推定量の期待値が真の値（母数）と一致する。

【直感的な意味】　何度も標本を取り出して推定を繰り返したとき、推定値の平均が真の値に一致するということ

標本平均は不偏推定量である

既に学んだ通り、標本平均 $\bar{X}$ は母平均 m の不偏推定量。

$$\mathrm{E}(\bar{X}) = m$$

不偏分散は不偏推定量である

Q2201 で証明したように、不偏分散は名前の通り、母分散 σ^2 の不偏推定量である。

$$s^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$$

$$\mathrm{E}(s^2) = \sigma^2$$

標本標準偏差は「不偏」なのか？

s^2 は不偏推定量だから「**不偏分散**」と呼ばれる。
では、その平方根である標本標準偏差

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum(x_i - \bar{x})^2}$$

は「**不偏標準偏差**」と呼んでいいのだろうか？

不偏性があるのなら、$\mathrm{E}(s) = \sigma$ が成り立つはず。

少し簡単な例で考えてみよう。

例　2 セットの標本から 2 つの不偏分散 s_1^2 と s_2^2 が得られたとする。この 2 つの平均は

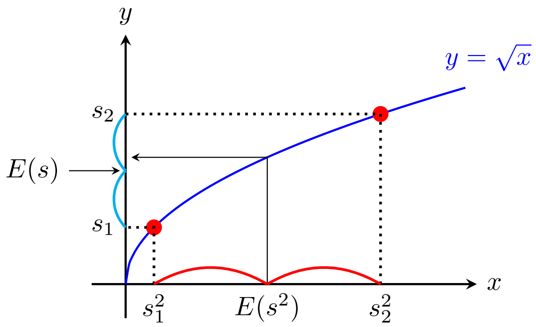
$$E(s^2) = \frac{s_1^2 + s_2^2}{2}$$

不偏性とはこの標本の数が増えていったときに $\mathrm{E}(s^2) = \sigma^2$ であるということすなわち、$\sqrt{\mathrm{E}(s^2)} = \sigma$

一方、不偏分散の平方根 s_1 と s_2 について平均を考えると

$$E(s) = \frac{s_1 + s_2}{2} \neq \sqrt{\frac{s_1^2 + s_2^2}{2}} = \sqrt{\mathrm{E}(s^2)}$$

よって、$E(s) \neq \sigma$ であるから不偏性は無い。



グラフから明らかに、

$$E(s) < \sqrt{E(s^2)} = \sigma$$

しかし、近い値であることには違いない。また、標本数が大きくなったときには、標本平均は母平均に近づくはずなので、毎回の不偏分散の差は小さくなるため、両者の値は極めて近い値になるだろう。

結論

$$\mathrm{E}(s) < \sigma$$

標本標準偏差 s は母標準偏差 σ の不偏推定量**ではない**。わずかに過小評価する傾向がある。

だから、「不偏標準偏差」ではなく「**標本標準偏差**」と呼ぶのが正確である。

実用上は問題ないのか？

標本標準偏差 s は不偏推定量ではないが、母平均の推定に用いても、実用上は問題ない。

理由：

- バイアス（偏り）は小さい
- 標本サイズ n が大きくなれば、バイアスはさらに小さくなる
- **t 分布**を使えば、このバイアスを考慮した正確な推定ができる

この後に続く学習「t 分布」では、標本標準偏差 s を使った正確な母平均の推定方法を学ぶ。t 分布は、s が不偏推定量でないことを自動的に補正してくれる。

問 1　次のうち不偏推定量であるものを記号で答えなさい。

(1) 標本平均 $\bar{x}$
(2) 不偏分散 s^2
(3) 標本標準偏差 s

答

++*+*+*+*+*+*+ **【解答】** *+*+*+*+*+*+*+*+

問 1　次のうち不偏推定量であるものを記号で答えなさい。

(1) 標本平均 $\bar{x}$

(2) 不偏分散 s^2

(3) 標本標準偏差 s

$\boxed{\text{答}}$　$\bar{x}$,　s^2