

確率統計

第15～16講

1

森岡 勇介

第15講

- 母分散がわかっている正規母集団について，母平均を区間推定する方法
- 一つが標本平均を使った母平均の区間推定（大標本の推定）
- もう一つは，分布についての知識を仮定しない「ノンパラメトリック」という方法（この本では紹介しない）

大標本の推定

母標準偏差がわかっている場合.

母標準偏差 σ と標本数 n を利用して, 標準偏差は
 $\frac{\sigma}{\sqrt{n}}$ で求めることができる.

大標本の推定

$\bar{x}$ を標本平均, μ が母平均, 分母に先ほどの標準偏差を代入すると, **95%信頼区間の数式**を立てることができ, これに当てはまる μ を母平均の95%信頼区間となる.

$$-1.96 \leq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq 1.96$$

まとめ

- 正規母集団で母標準偏差が σ とわかっている場合、母平均 μ を n 個の標本から推定するには、標本平均 $\bar{x}$ を計算して、

$$-1.96 \leq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq 1.96 \quad \text{を満たす}\mu\text{を残せばよい.}$$

上の式を変形して

$$\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}$$

これが母平均の95%信頼区間となる.

第16講

- 標本分散の話
- カイ二乗分布の登場
- 「標準正規分布する母集団から出てきた n 個のデータの2乗の和」という統計量を分析すること
- この時取り出したデータの数 n 個から自由度 n のカイ二乗分布という

標本分散

- ➡ STEP 1 標本平均を計算する.
- ➡ STEP 2 次に各標本から標本平均を引いて偏差をつくる.
- ➡ STEP 3 各偏差を2乗して合計し, 標本で割り算する.

$$(\text{標準分散 } s^2) = \{(\text{偏差 } 1)^2 + (\text{偏差 } 2)^2 + (\text{偏差 } 3)^2 \cdots\} / n$$

カイ二乗分布

標準正規分布（平均が0, S.D.が1の正規分布）であるような

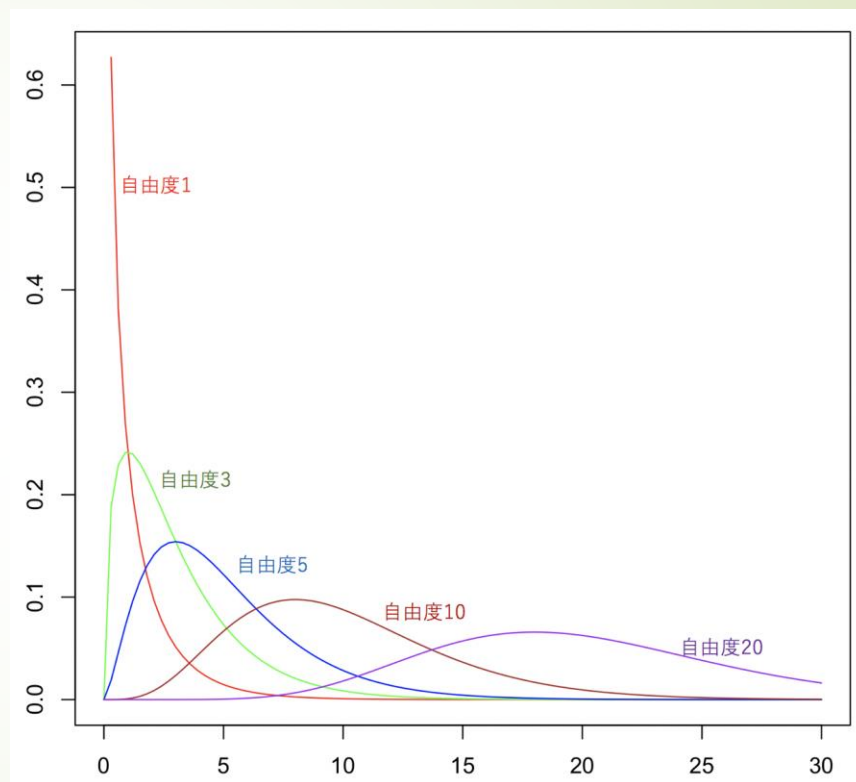
標準正規母集団から3個のデータを観測した場合を考えると,

観測データ x_1, x_2, x_3 に対して

$$V = x_1^2 + x_2^2 + x_3^2$$

と計算して数値Vを作ります.

その統計量をグラフにすると右のグラフの緑の線になります.



カイ二乗分布

標準正規分布の母集団からデータを三回観測する。「観測された3つの数値の二乗の和は3以上6未満である」と予言した場合、予言の的中率はどのくらいだろうか。

3以上である確率は表より

0.3916

6以上である確率は

0.1116

よって3以上6未満という予言は

$$0.3916 - 0.1116 = 0.28$$

よって28%の確率であたる。

自由度3のカイ二乗分布

x	X以上のVが観測される相対度数
0	1
1	0.8012
2	0.5724
3	0.3916
4	0.2614
5	0.1717
6	0.1116
7	0.0718
8	0.0460
9	0.0292
10	0.0185

まとめ

- 標準正規母集団からの n 個の標本に対してそれらを二乗して合計し、統計量 V を作ると、 V は自由度 n のカイ二乗分布をする。
- カイ二乗分布をする V は、二乗しているため 0 以上の値しかでてこない。また、0 に近いところの相対度数が大きく、0 から離れる数値の相対度数は急激に小さくなる。