統計学 ch9_解答

今日やること	母平均の区間推定について学びます。 区間推定で用いる信頼係数と信頼区間の意味を理解します。 母平均の区間推定に関する計算を行います。

pp.105　前回までの復習

正規分布	$Z=\dfrac{\bar{X}-\mu_X}{\sqrt{\dfrac{\sigma_X^2}{n}}}\sim N(0,1)$
カイ二乗分布	$U=\dfrac{\sum(X_i-\bar{X})^2}{\sigma_X^2}\sim \chi^2(n-1)$
t分布	$T=\dfrac{\bar{X}-\mu_X}{\sqrt{\dfrac{\widehat{\sigma_X^2}}{n}}}\sim t(n-1)$

※　色つきの部分は確率変数（可変）

※　それ以外の部分は固定値

pp.106　母平均の区間推定

区間推定　推定量の確率分布における区間を用いて母数を推定する

ある区間内に母数が含まれることを　信頼度　で示す

推定した区間は　信頼区間　で表す

信頼係数　（confedence coefficient）

信頼区間　（confedence interval）

信頼係数　$100(1-\alpha)$　%の信頼区間のことを

$100(1-\alpha)$ %信頼区間ともいう

（　下限値　,　上限値　）で表す

pp.109　母平均の95%信頼区間

ランダムな標本抽出を100回繰り返し行って

信頼区間を100回計算するとき

区間内に母平均を含むものは

100回のうち95回程度になるような区間

標本平均、t分布の%点、標本不偏分散を使って計算する

超重要

$$\left(\bar{x}-t_{0.025}(n-1)\sqrt{\frac{\widehat{\sigma_X^2}}{n}},\bar{x}+t_{0.025}(n-1)\sqrt{\frac{\widehat{\sigma_X^2}}{n}}\right)$$

pp.111　母平均の区間推定に関する計算

例題9-1　ポイントカード所有者2万人の母集団から50人を無作為抽出

標本数　$n = 50$

標本平均　$\bar{x} = 1.98$

標本分散　$S_x^2 = 2.6196$

母平均の95%信頼区間を求める

① 標本不偏分散を求める

$$\widehat{\sigma_X^2} = \frac{1}{n-1}\sum(x_i-\bar{x})^2 = \frac{n}{n-1}\frac{1}{n}\sum(x_i-\bar{x})^2 = \frac{n}{n-1}S_x^2$$

（pp.98 定義16 参照）より

$$\widehat{\sigma_X^2} = \frac{n}{n-1}S_x^2 = \frac{50}{49} \times 2.6196 \approx 2.6731$$

② t分布表から、t分布の%点を求める

今回は母平均の95%信頼区間なので

2.5%点 を求める $\left(\frac{1-0.95}{2}\right) = \left(\frac{0.05}{2}\right)$

※　この例の場合は表中に数字がないが、

$t_{0.025}(49) = 2.0096$　として計算することとする

③ t分布の%点、標本不偏分散、標本数を代入する

すなわち

$$t_{0.025}(49) \times \sqrt{\frac{\widehat{\sigma_X^2}}{n}} = 2.0096 \times \sqrt{\frac{2.6731}{50}} \approx 0.46466$$

④ 標本平均から±して上限値と下限値を求める

下限値

$$\bar{x} - t_{0.025}(n-1)\sqrt{\frac{\widehat{\sigma_X^2}}{n}} = 1.98 - 0.46466 \approx 1.54$$

上限値

$$\bar{x} + t_{0.025}(n-1)\sqrt{\frac{\widehat{\sigma_X^2}}{n}} = 1.98 + 0.46466 \approx 2.44$$

以上より、母平均の95%信頼区間は　$(1.54\,,2.44)$

pp.112　統計データからの区間推定

例題9-2　ポイントカード所有者2万人の母集団から7人を無作為抽出

標本平均、標本分散等は不明

得られた統計データは以下の表の通り

i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	3	0	0
2	0	-3	9
3	2	-1	1
4	7	4	16
5	5	2	4
6	3	0	0
7	1	-2	4
合計	21	0	34
平均値	3		

① 標本平均が不明の場合、

まず標本平均を計算

※ 有効桁数は多めにとる

② 平均値からの偏差を計算

※ 合計がゼロになる確認

③ 偏差平方和を計算

（統計学Ａの復習）

④ t分布表から2.5%点を求める

$t_{0.025}(7-1) =$ $\boxed{2.4469}$

⑤ 標本不偏分散を求める

$$\widehat{\sigma_X^2} = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = \frac{1}{7-1} \times 34 \approx 5.6667$$

⑥ t分布の%点、標本不偏分散、標本数を代入する

$$t_{0.025}(n-1) \times \sqrt{\frac{\widehat{\sigma_X^2}}{n}} = 2.4469 \times \sqrt{\frac{5.6667}{7}} \approx 2.2016$$

④ 標本平均から±して上限値と下限値を求める

下限値

$$\bar{x} - t_{0.025}(n-1)\sqrt{\frac{\widehat{\sigma_X^2}}{n}} = 3.0 - 2.2016 \approx 0.80$$

上限値

$$\bar{x} + t_{0.025}(n-1)\sqrt{\frac{\widehat{\sigma_X^2}}{n}} = 3.0 + 2.2016 \approx 5.20$$

以上より、母平均の95%信頼区間は $\boxed{(0.80\,, 5.20)}$

【練習問題】　下表は、あるクラスの英語の試験の受験者から抽出した15人分の得点である。

受験番号	得点	偏差	偏差平方和
1	28	-17.2	295.84
2	38	-7.2	51.84
3	32	-13.2	174.24
4	41	-4.2	17.64
5	42	-3.2	10.24
6	50	4.8	23.04
7	42	-3.2	10.24
8	68	22.8	519.84
9	52	6.8	46.24
10	28	-17.2	295.84
11	67	21.8	475.24
12	22	-23.2	538.24
13	27	-18.2	331.24
14	79	33.8	1142.44
15	62	16.8	282.24
合計	678		4214.4

Memo

① 標本平均と標本分散を求めよ

標本平均	45.2
標本分散	280.96

② 標本平均は正規分布にしたがい
標本分散の値を母分散として
考えることができるとした場合、
95%信頼区間を求めよ。

ヒント　標準正規分布表

$n =$ 15

$\bar{x} =$ 45.2

$\Pr(-1.96 < Z < 1.96) = 0.95$

$\widehat{\sigma_X^2} = S_x^2 =$ 280.96

(36.72 , 53.68)

③ 標本平均は正規分布にしたがい
標本分散は母集団の分散の
推定値に過ぎないと考える場合、
95%信頼区間を求めよ。

ヒント　t分布表

②の例では $\widehat{\sigma_X^2} = S_x^2$ 　だが
「推定値に過ぎない」と
考える場合は
「標本不偏分散」を求める

$\widehat{\sigma_X^2} =$ 301.0286

t%点は 2.1448

(35.59 , 54.81)

【練習問題】　下表は、ある日のコンビニ20店舗の売上高である。

店舗ＩＤ	売上高	偏差	偏差平方和
1	41	13.6	184.96
2	17	-10.4	108.16
3	14	-13.4	179.56
4	23	-4.4	19.36
5	33	5.6	31.36
6	38	10.6	112.36
7	36	8.6	73.96
8	26	-1.4	1.96
9	32	4.6	21.16
10	27	-0.4	0.16
11	41	13.6	184.96
12	25	-2.4	5.76
13	31	3.6	12.96
14	21	-6.4	40.96
15	27	-0.4	0.16
16	35	7.6	57.76
17	18	-9.4	88.36
18	19	-8.4	70.56
19	16	-11.4	129.96
20	28	0.6	0.36
合計	548		1324.8

Memo

① 標本平均と標本分散を求めよ

標本平均	27.4
標本分散	66.24

② 標本平均は正規分布にしたがい
標本分散の値を母分散として
考えることができるとした場合、
95%信頼区間を求めよ。

ヒント　標準正規分布表

$n =$ 20

$\bar{x} =$ 27.4

$\Pr(-1.96 < Z < 1.96) = 0.95$

$\widehat{\sigma_X^2} = S_x^2 =$ 66.24

(23.83 , 30.97)

③ 標本平均は正規分布にしたがい
標本分散は母集団の分散の
推定値に過ぎないと考える場合、
95%信頼区間を求めよ。

ヒント　t分布表

②の例では $\widehat{\sigma_X^2} = S_x^2$ 　だが
「推定値に過ぎない」と
考える場合は
「標本不偏分散」を求める

$\widehat{\sigma_X^2} =$ 69.7263

t%点は 2.093

(23.49 , 31.31)

今日の講義のまとめ

母平均の区間推定

母平均を含む統計量に基づいて行う

信頼係数 confidence coefficient

信頼区間 confidence interval

（下限 , 上限）

95%信頼区間

信頼係数95%の信頼区間を、95%信頼区間ともいう

ランダムな標本抽出を100回繰り返し行って

信頼区間を100回計算するとき

区間内に母平均を含むものは100回中95回程度になる区間

$$\left(\bar{x} - t_{0.025}(n-1)\sqrt{\frac{\widehat{\sigma_X^2}}{n}}, \bar{x} + t_{0.025}(n-1)\sqrt{\frac{\widehat{\sigma_X^2}}{n}}\right)$$

区間推定の計算

使用するもの

標本平均 $\bar{x}$

標本不偏分散 $\widehat{\sigma_X^2}$

t 分布のパーセント点 $t_{0.025}(n-1)$

$$\text{標本平均} \pm t\text{分布のパーセント点}\sqrt{\frac{\text{標本不偏分散}}{\text{標本数}}}$$

Memo