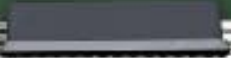





確率と統計

中山クラス
第16週(火曜クラス)



今日の予定

- ◆ 達成度確認試験の解説
第5回レポート解説
- ◆ 成績の集計, 評価の方法など
- ◆ 自己点検授業
- ◆ 授業評価アンケート

達成度確認試験の解説

配点：正解(○)：1点／設問

解答が正解に近い場合(△)：0.5点／設問

合計点：31点(答案用紙に記載)

→ 成績評価の際に40点満点に換算

問題 I (10点満点)

次の文章の空欄に下欄から適当な語句を選択せよ。答案用紙に番号を記入せよ。

「推測統計では、非常に大規模なデータ全体の統計的性質について、その一部を取り出したデータから推測することが行われる。データ全体を(ア)、一部を取り出したデータを(イ)、取り出すことを(ウ)という。(ア)の統計量を(エ)という。標本データから計算される統計量を(オ)という。(ア)の平均は(カ)、(イ)の平均は(キ)という。ある母数を推定するために用いられる標本統計量を(ク)、その値を(ケ)という。母数の値と(ケ)のずれは標本抽出に伴う誤差であり(コ)と呼ばれる。」

<選択肢>

1. 推定量, 2. 母集団, 3. 標本抽出, 4. 標本分布, 5. 母数, 6. 標本平均, 7. 推定値, 8. 標本統計量, 9. 母平均, 10. 標本誤差, 11. 標本

(ア)2, (イ)11, (ウ)3, (エ)5, (オ)8
(カ)9, (キ)6, (ク)1, (ケ)7, (コ)10

問題Ⅱ (3点満点)

正規母集団の母平均の推定に関して以下の問いに答えよ.

1. 正規分布 $N(\mu, \sigma^2)$ に従う母集団から n 個の標本を無作為に抽出したときの標本分布を求めよ.

$$N(\mu, \sigma^2/n)$$

2. 標準誤差を求めよ.

$$\sigma/\sqrt{n}$$

3. 標本平均は母平均の推定量であるが, その精度を高めるにはどのようにしたらよいか述べよ.

精度を上げるためには, 標準誤差 σ/\sqrt{n} を小さくする必要がある. σ は母数であるから, 調整できない. このため, 精度を上げるためには, サンプルサイズ n を大きくする.

問題Ⅲ(7点満点)

数学の得点は平均が7の正規分布に従うことが知られている. 次に示す10名の点数はこの母集団からの無作為抽出と考えると良いか検定せよ.

数学の点数: 5, 7, 4, 8, 6, 5, 9, 7, 6, 8

1. 帰無仮説と対立仮説を求めよ.

帰無仮説: 10名の数学の点数は平均が7の正規母集団からの無作為抽出である. (無作為抽出した正規母集団の平均は7である)

対立仮説: 10名の数学の点数は平均が7の正規母集団からの無作為抽出ではない. (~同文~平均は7ではない)

2. 片側検定か両側検定かを理由を付して述べよ.

平均が7より大きくても, 小さくても棄却されるので両側検定である.

3. 検定統計量 t を求めよ(式で示せ). 但し, 標本平均を \bar{X} , 不偏分散を $\hat{\sigma}^2$ とする.

分散が不明であるので, 次の検定統計量を用いる.

$$t = \frac{\bar{X} - \mu}{\hat{\sigma}/\sqrt{n}}$$

4. 検定統計量 t の実現値を求めよ.
但し, $\hat{\sigma} = 1.5811$, $\sqrt{10} = 3.1623$ とする.

与えられた標本データから \bar{X} を手計算で求める.

$$\bar{X} = 6.5$$

次に, 3の式を計算する.

$$t = -1.0$$

5. 検定統計量 t が従う確率分布を求めよ.
自由度が $n - 1 = 9$ の t 分布

6. 有意水準5%に対する棄却域を求めよ. 但し, 下記の「表 関数とその値」を参考にする.

両側検定であるから, 下側確率=2.5%, 上側確率=2.5%とする. 検定統計量が自由度=9のt分布に従うので, 下側確率=2.5%に対する境界値は表より $qt(0.025, 9) = -2.262157$ となる.

以上より, 棄却域は次のようになる.

$$t < -2.262157, \quad 2.262157 < t$$

7. 検定結果を理由を付して示せ. (帰無仮説が棄却される/されない. その結果, ○○○○であるとは言える/言えない)

検定統計量の実現値 $t = -1.0$ は棄却域に含まれないので, 帰無仮説は棄却されない. 従って, 10名の数学の得点は, 有意水準5%において平均=7の正規母集団からの無作為抽出である.

表 関数とその値

関数	p	df	関数の値	関数	p	df	関数の値
qt(p,df)	0.025	8	-2.306004	qchisq(p,df)	0.95	1	3.841459
		9	-2.262157			2	5.991465
		10	-2.228139			3	7.814728
	0.05	8	-1.859548		0.05	1	0.00393214
		9	-1.833113			2	0.1025866
		10	-1.812461			3	0.3518463

問題IV(7点満点)

以下のクロス集計表に関して、数学の好き嫌いと英語の好き嫌いの連関(or独立性)を検定せよ.

		英語		計
		好き	嫌い	
数学	好き	7	5	12
	嫌い	4	4	8
計		11	9	20

1. 帰無仮説と対立仮説を求めよ.

帰無仮説: 数学の好き/嫌い と、英語の好き/嫌い の間には連関がない(独立である).

対立仮説: 数学の好き/嫌い と、英語の好き/嫌い の間には連関がある(独立でない).

2. 片側検定か両側検定かを理由を付して述べよ.

独立性の検定にはカイ二乗検定を用いる. カイ二乗は正の値を取り, 連関が低いと小さな値となり, 連関が高いと大きな値となる. 従って, 片側検定となる.

3. 検定統計量 X^2 を求めよ(式で示せ). 但し, 観測度数を O_i , 期待度数を E_i とする.

$$X^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \dots + \frac{(O_k - E_k)^2}{E_k}$$

4. 検定統計量 X^2 の実現値を求めよ.

観測度数を $O_{11} = 7, O_{12} = 5, O_{21} = 4, O_{22} = 4$ とする.
これらに対する期待度数を求める.

$$E_{11} = \frac{12 \times 11}{20} = 6.6, \quad E_{12} = \frac{12 \times 9}{20} = 5.4,$$
$$E_{21} = \frac{8 \times 11}{20} = 4.4, \quad E_{22} = \frac{8 \times 9}{20} = 3.6$$

これらを3の式に代入して X^2 を求める.

$$X^2 = 0.135$$

5. 検定統計量 X^2 が従う確率分布を求めよ.

自由度が(行の数 -1) \times (列の数 -1) $=(2-1)\times(2-1)=1$ である
カイ二乗分布

6. 有意水準5%に対する棄却域を求めよ.
但し, 問題Ⅲにある「表 関数とその値」を参考にする事.

片側検定であり, 境界値より大きい領域が棄却域であるから, 自由度=1のカイ二乗分布において上側5%が棄却域になる. 従って, 表より $qchisq(0.95, 1) = 3.841459$ が境界値になり, 棄却域は次のようになる.

$$3.841459 < X^2$$

7. 検定結果を理由を付して示せ. (帰無仮説が棄却される/されない. その結果, ○○○○であると言える/言えない)

検定統計量 $X^2 = 0.135$ は棄却域に含まれないから, 帰無仮説は採択される. すなわち, 数学の好き/嫌いと英語の好き/嫌いには5%の有意水準で連関がないといえる.

問題 V (4点満点)

次の関数で計算される(処理される)内容を述べよ.

`dnorm(x, mean, sd)`

平均 = mean, 標準偏差 = sd の正規分布(確率密度関数)を表す.
x は確率変数.

`rnorm(n, mean, sd)`

平均 = mean, 標準偏差 = sd の正規分布から n 個の乱数を抽出する.

`pnorm(q)`

標準正規分布において下側確率を求める.

標準正規分布において確率変数(検定統計量)が $Z < q$ の値をとる確率を求める.

`pt(q, df)`

t 分布において下側確率を求める.

自由度が df の t 分布において, 確率変数(検定統計量)が $t < q$ の値をとる確率を求める.

第5回レポートの解説

教科書にある父母と娘の身長を用いる他、自分でもいろいろなデータを収集して統計解析を行っていた。

父母と娘のデータにおいては、父－娘の身長には相関があるが、母－娘、父－母の間には相関は見られなかった。

<評価した点>

いろいろな角度から解析している。

自分でデータを収集し、工夫して解析している。

<評価>

A+ 課題を満たしている。加えて、多角的に解析している。

自分でデータを収集している。

A 課題を満たしている。

B 課題を半分程度満たしている。

総合成績の評価

レポート	20点満点(平均:17.29点, 86.4%)
小テスト	40点満点(平均:25.34点, 63.4%)
達成度確認試験	40点満点(平均:29.07点, 72.7%)
<合計点>	100点満点(平均:69.95点)

*「F」の学生は除く

平均点調整(難易度調整) 全員に+2.7点

合格点(60点)に僅かに及ばない学生への対応

出席状況, レポート提出状況が良い学生に再チャレンジの機会を与える. レポート提出により, 合格点を与える.

<成績評価>

90点～	S	
80点～	A	
70点～	B	
60点～	C	
60点未満	D(不合格)	
出席不良	F(不合格)	6回以上欠席

配付資料の見方

名列	学籍番号	クラス	氏名	出席	公欠	欠席	レポ1 点数	レポ2 点数	レポ3 点数	レポ4 点数	レポ5 点数	20 点満点	小 テスト	40 点満点	達 確試験	40 点満点	合計100点	難 易度調整	成 績評価
----	------	-----	----	----	----	----	-----------	-----------	-----------	-----------	-----------	-----------	----------	-----------	----------	-----------	--------	-----------	----------

<レポート採点基準>

各レポートを5点満点，合計25点満点→20点満点に換算

A+	5	B+	4	C	3
A	4.5	B	3.75	C-	2.5
A-	4.25	B-	3.5		

難易度調整(平均点調整) +2.7点

再チャレンジのレポート課題

(1) 相関を検定するためのデータを収集する.

＜例＞USDドル／円レートと株価
金沢の気温と東京の気温

(2) 無相関検定を行う. 有意水準は5%とする.

提出期限: 8月14日(火)17:00時

提出場所: 21号館4階のレポート受け

不合格者への来年度の履修案内

来年度以降「確率と統計」は開講されません。不合格者は以下の点に注意してください。

- (1) 来年度の前学期履修手続き時に、教務課から提示される 代替科目を確認して履修すること。
- (2) 「確率と統計」の再履修クラスが開講された場合には、再履修クラスを履修すること。

自己点検授業

1. 確率と統計の内容、行動目標の再確認

2. 第5回レポート, 答案, 成績配布

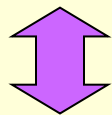
内容を確認し, 不明な点や修正等がある場合は授業時間中に申し出ること.

8月9日までに成績を登録します.

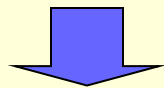
「確率と統計」の概要

学習教育目標: 確率や統計の基本概念を学び, 統計的資料の処理や解釈の方法を習得すること

多くの「確率と統計」の授業は数学的側面に重点を置く.



数学的には厳密だが, 実際に知識をデータに適用することが難しい.



数学的側面は最小限にし, コンピュータ演習を取り入れた. 統計処理のソフトウェアとして近年注目されているRの使用を基本にした.

行動目標

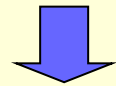
レポートおよび小テスト，達成度確認試験において，問題が解けた，あるいは内容を復習して理解したら，目標を達成したと考えられる。

- 統計解析システムRの基本操作ができる。
- 統計処理を計算機プログラムとして記述することができる。
- 離散的・連続的確率分布の例を挙げて，その数理的な意味を説明することができる。
- データの統計的処理の必要性を，基本統計量を算出して説明できる。
- 抽出された標本から母集団の特徴を推定するための適切な方法について説明できる。
- 適切な帰無仮説をたてて検定を実施することができる。

今後も

「確率と統計」は現実に良く使う、実用的な数学

- 創造実験, プロジェクトデザインなどで使う場面が多々ある.
- 現実社会でも統計の知識や技能が基礎になることは多い. 統計処理を行っていろいろな方針決定をしている.
→国の政策, 会社や部局の経営, 金融などなど



「使える統計」を目標に授業を実施した.

学業だけでなく, 生活にも活かしてくれるとうれしい.

授業評価アンケート

用紙を配布するので、今この場で記入する。
学生代表が収集して、教務課に提出する。

フィードバックが欲しいこと:

- 今後、他の科目のデータ処理でRを使うと思うか？
- Rの使い方で、もっと詳しい説明が必要だったか？
- 「確率と統計」の勉強にコンピュータ演習は不要か？