



確率と統計

中山クラス
第16週

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：34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

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

次の用語について説明せよ.

1. 標本分布

標本統計量に関する確率分布. 標本分布は度数分布ではなく, 確率分布である. 標本分布は母集団分布と標本統計量の種類とサンプルサイズが決まると理論的(数学的)に求まるものであり, 標本抽出されたデータから作成されるものではない.

2. 確率変数

ある事象における変数で, 値を確定できないが, その出現頻度や振る舞いは確率的表すことができる. 一例として, サイコロの目がある. n 回目に出るサイコロの目は確定できないが, 多数回振れば, 1の目は $1/6$ 回出ることが期待できる.

3. 単純無作為抽出

母集団からデータを抽出するときに、全てのデータが平等に選ばれる可能性がある標本抽出の方法.

4. 確率分布(離散変数, 連続変数)

ある確率変数がどのような値をどのような確率でとるかを表した分布. 確率変数が離散変数の場合は確率分布が確率そのものを表す. 連続変数の場合は, 確率そのものではなく, 確率密度を表す. 確率変数がある範囲の値をとる確率が面積で与えられる.

問題Ⅲ

数学の得点は $N(\mu, \sigma^2) = N(6, 4)$ 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て良いか検定せよ。

数学の点数: 5, 7, 6, 8, 4, 7, 5, 8, 6, 7

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 $N(\mu, \sigma^2) = N(6, 4)$ からの無作為抽出である。(または)母集団の平均は6である。

◇対立仮説: 無作為抽出ではない。(または)母集団の平均は6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。

母集団の平均が6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 Z を求めよ(式で示せ). 但し, 標本平均を \bar{X} , サンプルサイズを n とする.

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

4. 検定統計量の実現値を求めよ. 但し, $\sqrt{10} = 3.162$ とする.

手計算により, $\bar{X} = 6.3$ を求める. これを3の式に代入する.

$$Z = 0.474$$

5. 検定統計量が従う確率分布を求めよ.

平均=0, 分散=1である(標準)正規分布: $N(0,1)$

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

両側検定であるから, 5%を下側2.5%, 上側2.5%に分ける. 標準正規分布 $N(0,1)$ で確率が2.5%となる下側の検定統計量 Z の値は $qnorm(0.025) = -1.959964$ で与えられる. 確率分布は左右対称であるから上側の値は1.959964となる. 従って, 棄却域は次のようになる.

$$Z < -1.959964, \quad 1.959964 < Z$$

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が, $-1.959964 < Z = 0.474 < 1.959964$ となり, 棄却域にないので帰無仮説は棄却されない(採択される). 従って, 10名の数学の点数は $N(6, 4)$ に従う母集団からの無作為抽出であるといえる.

問題IV

以下に示す数学の得点と英語の得点の間に相関があるかを無相関検定により調べよ.

学生ID	1	2	3	4	5	6	7	8	9	10
数学得点	5	7	6	4	8	6	9	5	7	6
英語得点	6	7	5	6	9	7	8	4	8	7

1. 帰無仮説と対立仮説を求めよ.

帰無仮説: 数学と英語の得点には相関がない(相関係数=0)

対立仮説: 数学と英語の得点には相関がある(相関係数 $\neq 0$)

2. 片側検定か両側検定かを理由を付して述べよ.

相関係数がプラス側, マイナス側のいずれにずれても(正負の)相関が発生し, 帰無仮説は棄却されるので, 両側検定である.

3. 検定統計量 t を求めよ(式で示せ). 但し, 標本相関係数を r , サンプルサイズを n とする.

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

4. 検定統計量 t の実現値を求めよ. 但し, $r = 0.7413$, $\sqrt{1-r^2} = 0.6712$, $\sqrt{8} = 2.828$ とする.

上記の数値を3の式に代入して計算する.

$$t = 3.12$$

5. 検定統計量 t が従う確率分布を求めよ.

自由度が $n - 2 = 8$ の t 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.

両側検定であるから, 5%を下側2.5%, 上側2.5%に分ける. 自由度が8のt分布で確率が2.5%となる検定統計量 t の下側の値は $qt(0.025, 8) = -2.306004$ で与えられる. 確率分布は左右対称であるから, 上側の値は2.306004である. 従って, 棄却域は次のようになる.

$$t < -2.306004, \quad 2.306004 < t$$

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, ○○○○であると言える/言えない)

検定統計量の実現値 $t = 3.12 > 2.306004$ であるから棄却域にあり, 帰無仮説は棄却される. 従って, 数学と英語の得点の間には5%の有意水準で相関があると言える(相関係数は5%水準で有意である).

問題 V

次の関数で計算される(処理される)内容を述べよ.

1. `dnorm(x, mean, sd)`

平均 = `mean`, 標準偏差 = `sd` の正規分布を求める関数(確率密度関数). `x` は確率変数.

2. `rnorm(n, mean, sd)`

平均 = `mean`, 標準偏差 = `sd` の正規分布に従う乱数を `n` 個発生する関数.

3. `pnorm(q)`

標準正規分布において, 確率変数(検定統計量)が $Z < q$ となる確率を求める関数.

4. `pt(q, df)`

自由度が df である t 分布において、確率変数(検定統計量)が $t < q$ となる確率を求める関数.

5. `qchisq(p, df)`

自由度が df のカイ二乗分布において、確率が $P(0 < X^2 < q) = p$ となる確率変数(検定統計量) q を求める関数.

6. `pchisq(q, df)`

自由度が df のカイ二乗分布において、確率変数(検定統計量)が $0 < X^2 < q$ となる確率を求める関数.

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:16点, 80%)
小テスト	40点満点(平均:22.34点, 55.85%)
達成度確認試験	40点満点(平均:29.39点, 73.48%)
<合計点>	100点満点(平均:67.74点)

*「F」の学生は除く

平均点調整(難易度調整) 全員に+5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	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	達成度確認試験	40 点満点	合計100点	難易度調整	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) +5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10日(金)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

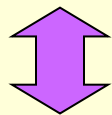
内容を確認し, 不明な点や修正等がある場合は授業時間中に申し出ること.

8月9日までに成績を登録します.

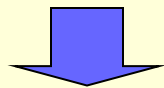
「確率と統計」の概要

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

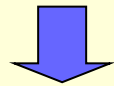
レポートおよび小テスト，達成度確認試験において，問題が解けた，あるいは内容を復習して理解したら，目標を達成したと考えられる。

- 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。

今後とも

「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？