

第6回数学科リレー講座 (初日)

Q&Aで知る本講座のガイダンス

川崎真澄

Contents

- 1日目 6日間の概観(川崎)
- 2日目 平均, 確率 (縫部先生)
- 3日目 相関(平山先生)
- 4日目 いろいろな分布(柴山先生)
- 5日目 推定(網谷先生)
- 6日目 検定(兼子先生)

本日のテーマ

平均 & 確率の基本を知ろう

Q1 テーマ:平均の盲点

「水深の平均が5cmの用水路で溺死者あり！」との一報あり。
この用水路でおぼれた原因を「統計的に考える」と、次のどれが濃厚といえるだろうか？

- ①酔っ払いが浅い用水路で転倒し、
起き上がれずに溺れた。
- ②乳幼児が溺れた。
- ③大人も完全に沈み込むような深いところには
まって溺れた。

A1. ③と考えるのが妥当である。
水深の平均が5cmというだけでは、ほとんどのところが水深2cmでしかないが、ある区間だけが水深2mであることも考えられるのだ。

平均は便利な目安。

されど

"平均では分からぬもの"

あり！

⇒そこで、

平均以外のデータの特徴づけを考えることが2日目以降のテーマである。これを頭に入れておいてください。

Q2. テーマ: 確率の基本

サイコロを2つ使ってそれらの目の和が偶数(丁)か奇数(半)かに張る「丁半ばくち」で“イカサマ”をしていると専らの噂の賭場あり。客に変装して、じっとその様子をうかがっていた正義の味方。やおら立ち上がって曰く、

「おう！片方のサイコロは1から6まで
でているようだがこのサイコロはさっ
きから1しかでていないぜ。イカサマ
のサイコロってわけだな！貸してみ
ろ！！」

ととりあげざまにガリッと噛んだら、
なんと鉛がポロッと落ちてきた。

「この悪党が！こうしてくれるわ！」
と言うが早いとそのサイコロふりを
叩き切ってしまった。

そしてこう言った

「自業自得だぜ、悪党め」。

湧きあがる大歓声！！

さて、この正義の味方は、

- ①叩き切ったことに賛否はあろうが正義をいかんなく発揮した。
- ②確かにイカサマであったろうが叩き切ったのはやりすぎ。
- ③この正義の味方は即刻裁判にかけて有罪にすべき。

Q2. ③

+	1	2	3	4	5	6
1	②	3	④	5	⑥	7

○は丁。

つまり、鉛入りのサイコロであろうと、
それが片方だけなら丁と半がでる確
率は半々ずつ！ということは、
イカサマにはなっていないのだ！

Q3. テーマ: 確率のゆらぎ

ズバリ, 統計学的に世の中には運の
良い人と, 悪い人が

- ① 存在する
- ② 存在しない

Q3. ①

もし、人生に無限に近い長さがあれば、
誰もが幸運、不運は同等に現れると
考えてよいが、

人生には限りがある！

のであった。

ゆえに、①と考えられる。

Q4. テーマ:期待値(これも平均)

国公認のギャンブルである宝くじと競馬。君が成人後, どちらかを投資として選ばねばならなくなったとしよう。さて, 有利なのはどちらだろうか?

A4. 競馬

宝くじの平均当選金は40%。

競馬の平均当選金は75%。

例えば、平均当選金が40%のくじとしては、次のようなものも考えられる。君には魅力的に映るだろうか？

(Q) 1本100円で50本売り出す。
このくじの平均当選金は？

等	当選金	本数
1	300円	2本
2	200円	4本
3	100円	6本
外れ	0円	38本

(A) 40円

それでも夢は買いたい・・・かな？！

Q5. テーマ:標準偏差(重要!!)

ある駅前発の午前7時半に到着する予定のA社とB社が運行するバスのある1週間にわたる到着時間は次の通りである:

	月	火	水	木	金	土	日	平均
A社	37	33	29	30	27	32	29	31
B社	28	32	30	33	30	29	35	31

分かることは、A社のバスもB社のバスも共に、先々週も先週も、平均して1分遅れての到着ということである。

つまり、到着時間の平均を見る限り、両社の差はないが、より正確に運行しているのはA社でとB社のどちらであろうか？

A5. B社である

まず、両社のバスの運行の差を見るために、到着時間の平均からの“振れ幅”を見てみよう。振れ幅の算出方法として、

(到着時間－到着時間の平均)の平均を考える人は少なくないと思われる。

ところが、これは無意味である。

実際、計算してみると、

A社:

$$\{(37-31)+(33-31)+(29-31)+(30-31) \\ + (27-31)+(32-31)+(29-31)\} \div 7 = 0$$

B社:

$$\{(28-31)+(32-31)+(30-31)+(33-31) \\ + (30-31)+(29-31)+(35-31)\} \div 7 = 0$$

となり, どちらも0となるからである。

果たして, これは偶然か?

答えは、常に0である。

例えば、簡単のために、3日間で考えてみよう。

	1日目	2日目	3日目	平均
到着時刻	a	b	c	d

とする。

このとき、 $(a+b+c) \div 3 = d$ つまり、 $a+b+c=3d$ なのだから、

$$(a-d)+(b-d)+(c-d) \div 3 = \{(a+b+c)-3d\} \div 3 \\ = 0$$

というわけである。

これは4日以上にしても同様である。

よって、他の算出方法を考えなければならない。

さきの方法だと、例えば、

7時28分着だと、平均から-3分

7時31分着だと、平均から+3分

であるが、

“振れ幅”

という意味では、

“共に平均から3分の違い”

といえる。

そこで、まず、 $+$ 、 $-$ を相殺するために、

(到着時間 $-$ 到着時間の平均)の
“二乗”の平均

をとってみよう:

A社:

$$\{(37-31)^2 + (33-31)^2 + (29-31)^2 + (30-31)^2 + (27-31)^2 + (32-31)^2 + (29-31)^2\} \div 7 \doteq 9.4$$

B社:

$$\{(28-31)^2 + (32-31)^2 + (30-31)^2 + (33-31)^2 + (30-31)^2 + (29-31)^2 + (35-31)^2\} \div 7 \doteq 5.1$$

となる。

ただ、これだと、単位が変わってしまう。

つまり、(分)ではなく、(分) である。

これを解消するために、

正の平方根をとる

ことにしよう。

すると,

A社の到着時間は平均7時31分であるが, その
振れ幅は,

およそ3分4秒

B社の到着時間も平均7時31分であるが, その
振れ幅は,

およそ2分15秒

であることが分かる。

すなわち、到着時間の平均は同じ
であっても、B社のバスの方がより
正確な運行をしているということが
いえるのである。

ここで、

到着時間－到着時間の平均
を“偏差”といい、

偏差の二乗の総和の平均値を
“分散”という。また、

分散の正の平方根を
“標準偏差”という。

この標準偏差が

“統計量としては最も大事な数値”

の一つと言ってよく、その重要性を理解することが本年のリレー講座の目的とも言えます。

さきに、平均の深さが浅くても、部分的に深い水たまりがあることをみましたが、平均だけでは特徴づけが難しいデータを

“よく観察するための物差し”

の一つとなりうる量、それが標準偏差であるといえます。

データは実相に踏み込まずに、
“使用者に都合よく使う”
こともできます。
その実例が明日以降示される
ことでしょう。

さあ、君は果たして、都合よく使う側になるか？それとも、見破る側になるか？

残り5日間でよく勉強して、そしてよく考えてください。

参考文献

なお、以下の文献には大変お世話になりました。
皆さんにもご一読をお勧めします：

- ① 統計カクイズ 涌井良幸著
実務教育出版・刊
- ② 図解雑学 統計 今野紀雄著
ナツメ社・刊
- ③ 確率統計(なべつぐのあすなろ数学) 渡辺次男著
旺文社・刊