## 3日目

# 統計的な推測

イワナは渓流釣りの対象魚として人気があり，各地の河川でその保護や増殖の活動が行われています。
しかし，その活動の成果を調べるためにイワナの数を調査しようと思っても，水中を動き回るイワナを１尾ずつすべて数えることはできません。どのようにすれば，イワナの数を調査することができるのでしょうか。

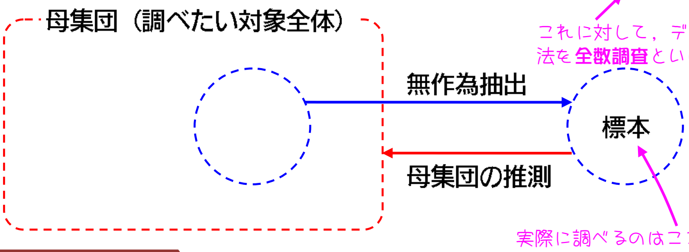
# 1　一部だけ調べろ!!　標本調査とは?

**ねらい**　標本調査とは何か，標本調査の手法について理解しよう。

## 1　標本調査とは?

大量のデータの情報を調べるとき，いちいち全部のデータ（**母集団**）を調べていたらキリがありません。そこで，実際には全部のデータの中から何個かをランダムに抽出し，抽出したデータ（**標本**）を調べてから母集団のデータの情報を予測するという手法を用います。この手法が**標本調査**です。

これに対して，データのすべてを調べる方法を**全数調査**といいます。




実際に調べるのはここです。

### 標本調査についての用語集

① 母集団の平均，分散，標準偏差を，それぞれ**母平均**，**母分散**，**母標準偏差**といいます。
② 標本のデータの個数を**標本の大きさ**といいます。また，標本の平均，分散，標準偏差を，それぞれ**標本平均**，**標本分散**，**標本標準偏差**といいます。

ここでは，これらの量の間にどんな関係があるか，下の例題を追っていく中で見出してみましょう。

**例題**

これが標本。

箱の中に大量のカードが入っており，中身の 3/4 には「2」，1/4 には「1」が書かれています。無作為に 2 枚のカードを取り出し，取り出したときに書かれているカードの数字を順に $X_1$，$X_2$ とします。

このとき，標本平均 $\overline{X}=\dfrac{X_1+X_2}{2}$ の平均 $E(\overline{X})$ および分散 $V(\overline{X})$ を求めてみましょう。

まず，$X_1$，$X_2$ のとり得る値とそのときの標本平均 $\overline{X}$，そのときの確率をそれぞれ求めてみると，右の表のようになります。よって，$\overline{X}$ の確率分布は

| $(X_1,\ X_2)$ | 標本平均 $\overline{X}$ | 確率 |
|---|---|---|
| (1，1) | 1 | $\frac{1}{4}\cdot\frac{1}{4}=\frac{1}{16}$ |
| (1，2) | $\frac{3}{2}$ | $\frac{1}{4}\cdot\frac{3}{4}=\frac{3}{16}$ |
| (2，1) | $\frac{3}{2}$ | $\frac{3}{4}\cdot\frac{1}{4}=\frac{3}{16}$ |
| (2，2) | 2 | $\frac{3}{4}\cdot\frac{3}{4}=\frac{9}{16}$ |

| $\overline{X}$ | 1 | $\frac{3}{2}$ | 2 | 計 |
|---|---|---|---|---|
| $P$ | $\frac{1}{16}$ | $\frac{6}{16}$ | $\frac{9}{16}$ | 1 |

となりますね。ってなワケで

平均　$E(\overline{X})=1\cdot\dfrac{1}{16}+\dfrac{3}{2}\cdot\dfrac{6}{16}+2\cdot\dfrac{9}{16}=\dfrac{7}{4}$

$E(\overline{X}^2)=1^2\cdot\dfrac{1}{16}+\left(\dfrac{3}{2}\right)^2\cdot\dfrac{6}{16}+2^2\cdot\dfrac{9}{16}=\dfrac{101}{32}$

分散　$V(\overline{X})=E(\overline{X}^2)-\{E(\overline{X})\}^2$

$=\dfrac{101}{32}-\left(\dfrac{7}{4}\right)^2=\dfrac{3}{32}$

いったん整理しておくと……
ここで求めたのは，
・**標本平均の平均（期待値）**
・**標本平均の分散**
だよ。

と求まります！

では，今の話と母平均，母分散との関係を見ていきます。

**例題**（再放送）

箱の中に大量のカードが入っており，中身の3/4には「2」，1/4には「1」が書かれています。
今度は，カードに書かれている数字の母平均 $m$，および母分散 $\sigma^2$ を求めてみましょう。

母集団は，箱の中のカードすべてなので，母平均は箱の中のすべてのカードの数字の平均値，母分散は箱の中にあるすべてのカードの数字の分散を求めれば OK。

母平均　$m=\frac{3}{4}\cdot 2+\frac{1}{4}\cdot 1=\frac{7}{4}$

母分散　$\sigma^2=1^2\cdot\frac{1}{4}+2^2\cdot\frac{3}{4}-\left(\frac{7}{4}\right)^2=\frac{3}{16}$

ここで求めたのは，
・母平均
・母分散
だよ。

と求まります。

さて，ここからが本題。

一致！

標本平均の平均：$\frac{7}{4}$　　母平均：$\frac{7}{4}$

標本平均の分散：$\frac{3}{32}$　　母分散：$\frac{3}{16}$

母分散を，標本の大きさで割ったものと一致！

ということで，ひょっとして

・(標本平均の平均)＝(母平均)

・(標本平均の分散)＝(母分散)÷(標本の大きさ)

が成り立つのではないか！？という気がしてきます。

っつーことで，証明してみましょう。　時間がなければ読み飛ばして OK！

母平均 $m$，母分散 $\sigma^2$ の母集団から大きさが $n$ の標本を選び，それぞれの標本の値が $X_1$，$X_2$，……，$X_n$ であるとします。これらの各変数は，大きさ 1 の標本の確率変数とみなされ，それぞれが母集団分布に従うから

$$E(X_i)=m,\quad V(X_i)=\sigma^2\quad(i=1,\ 2,\ \cdots\cdots,\ n)$$

が成立します。

少々わかりにくいのですが…
例えば，$X_1$ であれば，次の分布に従います。

| $X_1$ | 2 | 1 | 計 |
|---|---|---|---|
| $P$ | $\frac{3}{4}$ | $\frac{1}{4}$ | 1 |

他も同様です。母集団の分布と同じですね。

標本平均 $\overline{X}$ は

$$\begin{aligned}E(\overline{X})&=E\left(\frac{1}{n}(X_1+X_2+\cdots\cdots+X_n)\right)\\&=\frac{1}{n}\{E(X_1)+E(X_2)+\cdots\cdots+E(X_n)\}\\&=\frac{1}{n}(m+m+\cdots\cdots+m)\\&=\frac{1}{n}\cdot nm=m\end{aligned}$$

で，標本調査をする場合，母集団の大きさが非常に大きい場合を想定しているので，そこからいくつかのデータを無作為に選ぶ場合，それらの値は互いに独立に定まると仮定して OK です。

要は，1 枚ひいたとしても，（全体はあまりに多いから）全体の枚数は変わらないと考えてもいいよねってことです。

というわけで，標本平均の分散 $V(\overline{X})$ は

$$\begin{aligned}V(\overline{X})&=V\left(\frac{1}{n}(X_1+X_2+\cdots\cdots+X_n)\right)\\&=\frac{1}{n^2}\{V(X_1)+V(X_2)+\cdots\cdots+V(X_n)\}\\&=\frac{1}{n^2}(\sigma^2+\sigma^2+\cdots\cdots+\sigma^2)\\&=\frac{1}{n^2}\cdot n\sigma^2=\frac{\sigma^2}{n}\end{aligned}$$

となって成立しますね。めでたし，めでたし。

統計的な推測

**COLUMN：それぞれの量の意味**

結局のところ，何が求まったのか，どんな意味があるのか，ここで整理しておきましょう。具体的には，さっきの例で求まった

$$m=\frac{7}{4},\quad \sigma^2=\frac{3}{16},\quad E(\overline{X})=\frac{7}{4},\quad V(\overline{X})=\frac{3}{32}$$

の意味を考えてみましょう。

① **母平均 $m$ の意味**

箱の中のカードを **1 枚**選んで調べるということを，**すべてのカードに行う**と，そのデータの平均は $\frac{7}{4}$ になるよ，ってこと。← これが「大きさ 1 の標本の確率分布」＝「母集団分布」って意味です。

② **母分散 $\sigma^2$ の意味**

①と同様，箱の中のカードを 1 枚選んで調べるということを，すべてのカードに行うと，そのデータの分散は $\frac{3}{16}$ になるよ，ってこと。

③ **標本平均の平均 $E(\overline{X})$ の意味**

この例の場合，大きさ 2 の標本を考えているので，カードを無作為に 2 枚選んでその数字の平均を $\overline{X}$ としていて，$\overline{X}$ のとり得る値は，あるときは $\overline{X}=\frac{1+1}{2}=1$ となり，あるときは $\overline{X}=\frac{1+2}{2}=\frac{3}{2}$ となる。

で，$E(\overline{X})=\frac{7}{4}$ とは，「大きさ 2 の標本を選び標本平均 $\overline{X}$ を求める」ということを**何回も繰り返す**と，その平均 $E(\overline{X})$ はほぼ $\frac{7}{4}$ になるよ，ってこと。

④ **標本平均の分散 $V(\overline{X})$ の意味**

③と同様に，カードを無作為に 2 枚選んでその標本平均 $\overline{X}$ を求めるということを何回も繰り返すと，$\overline{X}$ の分散 $\sigma(\overline{X})^2$ がほぼ $\frac{3}{32}$ になるよ，ってこと。

大切なことなので，最後にまとめておきましょう。

**標本平均の公式**

① 任意の変量 $x$ について，母集団での母平均を $m$，標本平均を $\overline{X}$ とすると，$\overline{X}$ の平均 $E(\overline{X})$ について

$$E(\overline{X})=m$$

が成り立つ。

② 母集団の母分散を $\sigma^2$，$\overline{X}$ の分散を $V(\overline{X})$ とし，標本の大きさ $n$ に対して母集団の大きさが十分大きい（つまり，標本が互いに独立）とすると

$$V(\overline{X})=\frac{\sigma^2}{n}$$

が成り立つ。

## 2 中心極限定理

p.34～p.36 で，平均が $m$，分散が $\sigma^2$ の母集団から，大きさ $n$ の標本を無作為に選び標本平均を $\overline{X}$ とすると，

$$E(\overline{X})=m\,,\qquad V(\overline{X})=\frac{\sigma^2}{n}$$

が成立することを確認しました。

実は，標本の大きさ $n$ を大きくしていくと，$\overline{X}$ の確率分布は近似的に正規分布に従うということが知られています。

つまり，**母集団がどんな分布であっても，母集団から多くのデータを集めた標本の標本平均 $\overline{X}$ は正規分布に従う**ってことなんです！！ ← またもや正規分布表が活躍するわけですね。

これが**中心極限定理**と呼ばれる定理です。

**中心極限定理**

母平均 $m$，母分散 $\sigma^2$ の十分大きな母集団から大きさ $n$ の標本を無作為に選び標本平均を $\overline{X}$ とし，$n$ は十分に大きいとします。

このとき，$\overline{X}$ は平均が $m$，分散が $\dfrac{\sigma^2}{n}$ の正規分布 $\boldsymbol{N\left(m,\ \dfrac{\sigma^2}{n}\right)}$ に近似的に従う！！

例えば，母平均が 100，母分散が 1600 の母集団から大きさ 25 の標本を無作為に選び，その標本平均を $\overline{X}$ としたとき，$90\leqq\overline{X}\leqq110$ となる確率 $P(90\leqq\overline{X}\leqq110)$ のおよその値を求めてみましょう。その手順をぜひ覚えてくださいね。

① 標本の大きさ 25 を「十分大きい」とみなすと，$\overline{X}$ は正規分布 $N\left(100,\ \dfrac{1600}{25}\right)$ に近似的に従います。

② $\overline{X}$ の標準偏差は $\sqrt{\dfrac{1600}{25}}=\dfrac{40}{5}=8$

③ $P(90\leqq\overline{X}\leqq110)$ を求めるために

$$z=\frac{\overline{X}-(\overline{X}\text{の平均})}{(\overline{X}\text{の標準偏差})}=\frac{\overline{X}-100}{8}$$ ← 前回の話を思い出して！　標準化ってヤツです。

とおきます。$\overline{X}$ が正規分布に近似的に従うので $z$ は標準正規分布に近似的に従い，正規分布表が使えることになります。

④ $\overline{X}$ の範囲「$90\leqq\overline{X}\leqq110$」を $z$ の範囲に変換します。

$$90\leqq\overline{X}\leqq110$$
$$-10\leqq\overline{X}-100\leqq10$$
$$-\frac{5}{4}\leqq\frac{\overline{X}-100}{8}\leqq\frac{5}{4}$$
$$-1.25\leqq z\leqq1.25$$

⑤ あとは，$P(-1.25\leqq z\leqq1.25)$ を求めるだけです。

$$P(-1.25\leqq z\leqq1.25)=P(-1.25\leqq z\leqq0)+P(0\leqq z\leqq1.25)$$
$$=P(0\leqq z\leqq1.25)+P(0\leqq z\leqq1.25)$$ ← 正規分布表の使い方を思い出して！ 標準正規分布は $y$ 軸について対称です。
$$=0.3944\times2=0.7888$$

⑥ ということで，$P(90\leqq\overline{X}\leqq110)\fallingdotseq0.79$ とわかります！！

# 2 最重要！ 母平均・母比率の推定

ねらい 共通テスト（センター試験）で毎年のように問われています。絶対マスターを！！

## 1 信頼度とは？

標本調査では，母集団の一部のデータを抽出し，抽出したデータから母集団のデータを予測するため，正確なデータを求めることができません。

そこで，ある範囲内に正確なデータがある確率はどのくらいかを表したものを**信頼度**といい，ある範囲のことを**信頼区間**と呼びます。

**信頼度 95 %の信頼区間**：信頼区間内に正しいデータがある確率が 95 %

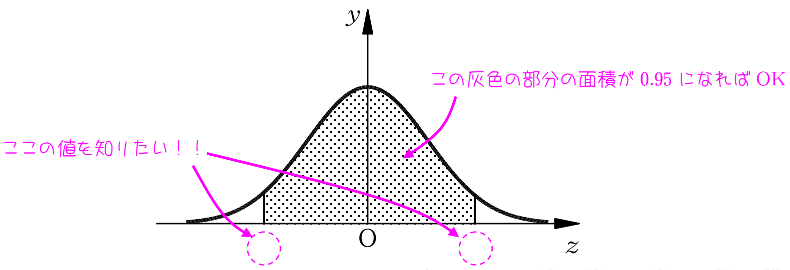
**信頼度 99 %の信頼区間**：信頼区間内に正しいデータがある確率が 99 %

みたいな言い方をします。

数 B「確率分布と統計的な推測」では，標本を正規分布とみなして標本から計算された平均，標準偏差から信頼区間を求めることになります。具体的には，平均が信頼区間の中央となり，そこから標準偏差と信頼度によって信頼区間を決定します。

標準正規分布において，信頼度 95 %の信頼区間がどの程度になるか確認してみましょう。

信頼度 95 %ということは，信頼区間内に正しいデータがある確率が 95 %ということです。これを図でイメージすると……




標準正規分布の分布曲線は $y$ 軸について対称なので，右半分の面積が $0.95 \div 2 = 0.475$ になる $z_0$ の値を正規分布表から読み取ればよいことになります！

右の表をよく見ると……

$$z_0 = 1.96$$

となることがわかります。これによって，信頼度 95 %に相当する信頼区間は，**平均±1.96** 程度の範囲ってことになりますね。

同様に，信頼度 99 %の信頼区間は

**平均±2.58**

ということもわかります。

| $z_0$ | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.0 |
|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0000 | 0.0040 | 0.0080 | 0.0120 | 0.0160 | 0.0199 | 0.0239 | 0.02 |
| 0.1 | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 | 0.0596 | 0.0636 | 0.06 |
| 0.2 | 0.0793 | 0.0832 | 0.0871 | 0.0910 | 0.0948 | 0.0987 | 0.1026 | 0.10 |
| 0.3 | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 | 0.1368 | 0.1406 | 0.14 |
| 0.4 | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.1700 | 0.1736 | 0.1772 | 0.18 |
| 0.5 | 0.1915 | 0.1950 | 0.1985 | 0.2019 | 0.2054 | 0.2088 | 0.2123 | 0.21 |
| 0.6 | 0.2257 | 0.2291 | 0.2324 | 0.2357 | 0.2389 | 0.2422 | 0.2454 | 0.24 |
| 0.7 | 0.2580 | 0.2611 | 0.2642 | 0.2673 | 0.2704 | 0.2734 | 0.2764 | 0.27 |
| 0.8 | 0.2881 | 0.2910 | 0.2939 | 0.2967 | 0.2995 | 0.3023 | 0.3051 | 0.30 |
| 0.9 | 0.3159 | 0.3186 | 0.3212 | 0.3238 | 0.3264 | 0.3289 | 0.3315 | 0.33 |
| 1.0 | 0.3413 | 0.3438 | 0.3461 | 0.3485 | 0.3508 | 0.3531 | 0.3554 | 0.35 |
| 1.1 | 0.3643 | 0.3665 | 0.3686 | 0.3708 | 0.3729 | 0.3749 | 0.3770 | 0.37 |
| 1.2 | 0.3849 | 0.3869 | 0.3888 | 0.3907 | 0.3925 | 0.3944 | 0.3962 | 0.39 |
| 1.3 | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 | 0.4115 | 0.4131 | 0.41 |
| 1.4 | 0.4192 | 0.4207 | 0.4222 | 0.4236 | 0.4251 | 0.4265 | 0.4279 | 0.42 |
| 1.5 | 0.4332 | 0.4345 | 0.4357 | 0.4370 | 0.4382 | 0.4394 | 0.4406 | 0.44 |
| 1.6 | 0.4452 | 0.4463 | 0.4474 | 0.4484 | 0.4495 | 0.4505 | 0.4515 | 0.45 |
| 1.7 | 0.4554 | 0.4564 | 0.4573 | 0.4582 | 0.4591 | 0.4599 | 0.4608 | 0.46 |
| 1.8 | 0.4641 | 0.4649 | 0.4656 | 0.4664 | 0.4671 | 0.4678 | 0.4686 | 0.46 |
| 1.9 | 0.4713 | 0.4719 | 0.4726 | 0.4732 | 0.4738 | 0.4744 | $\boxed{0.4750}$ | 0.47 |
| 2.0 | 0.4772 | 0.4778 | 0.4783 | 0.4788 | 0.4793 | 0.4798 | 0.4803 | 0.48 |

## 2 具体例で理解する母平均の推定

比較的簡単に求めることのできる標本平均 $\overline{X}$ と標本分散 $s$ から母平均 $m$ を推定する方法を，具体例を通じて理解してもらいます。

**例題**

ある洋菓子店では手作りで板チョコを大量に作っている。そのうちの $n=100$（枚）の重さを調べたところ，1枚の重さの標本平均は $\overline{X}=50$ (g)，標本標準偏差が $S=2$ (g) であった。

板チョコ1枚の重さの母平均を $m$ (g)，母標準偏差を $\sigma$ (g)とし，$m$ に対する信頼度 95 %の信頼区間を求めよう。

① 母分散は $\sigma^2$ であり，$n=100$ は十分大きいので $\overline{X}$ は正規分布 $N\left(m,\ \frac{\sigma^2}{100}\right)$ に近似的に従い，$\overline{X}$ の平均を $E(\overline{X})$，標準偏差を $s(\overline{X})$ とすると

$$E(\overline{X})=m,\qquad s(\overline{X})=\sqrt{\frac{\sigma^2}{100}}=\frac{\sigma}{10}$$

$\sigma$ は標本標準偏差 $S=2$ とほぼ等しいと考えて

母標準偏差がわかっていない場合は，標本標準偏差を代わりに使うのが一般的です。

$$s(\overline{X})=\frac{2}{10}=\frac{1}{5}$$

としてよい。

② 正規分布表を用いるために

$$z=\frac{\overline{X}-E(\overline{X})}{s(\overline{X})}=\frac{\overline{X}-m}{\frac{1}{5}}=5(\overline{X}-m)$$

とおきます。このことによって，$z$ は標準正規分布に近似的に従い，正規分布表が使えるようになりました。

③ 信頼度 95 %の信頼区間，つまり，$y$ 軸について対称な区間 $-z_0 \leqq z \leqq z_0$ を $z$ が満たす確率が 95 %（$=0.95$）となる $z_0$ は，さっきの議論から

$$z_0=1.96$$

④ $-1.96 \leqq z \leqq 1.96$ となる確率が 95 %となるので，これを $z=5(\overline{X}-m)$ を用いて，$m$ の範囲に書き直します。つまり

$$-1.96 \leqq 5(\overline{X}-m) \leqq 1.96$$
$$-1.96 \leqq 5(50-m) \leqq 1.96$$
$$-0.392 \leqq 50-m \leqq 0.392$$

整理して，小数第3位を四捨五入すると

$$\mathbf{49.61 \leqq m \leqq 50.39}$$

となります！！

このことが言っているのは，「板チョコ1枚の重さの母平均 $m$ (g)が確率 95 %でこの範囲に入っている」ってことです。言い換えると，「**上のように信頼区間を求めることを何回も繰り返せば，そのうちの 95 %ぐらいでその信頼区間に母平均 $m$ が入る**」となります。

板チョコすべての重さを調べなくても，その一部を調べただけでこれだけわかるというのが，統計の面白いところですね。

## 3 母比率の推定

いよいよ最終項目です。ここでは，「母集団の中である事象が起こる確率」を，比較的簡単に計算ができる「標本の中である事象が起こる確率」から推定する方法を伝授します。

母集団の中で，ある事象 $A$ を満たす確率のことを**母比率**と呼び，標本の中で事象 $A$ を満たす確率のことを**標本比率**といいます。

母集団全体を調べるのは通常非常に困難なので，母比率 $p$ はわからないことがほとんどです。

そこで，標本の大きさ $n$ が十分に大きいときは，母比率 $p$ は標本比率 $R$ にほぼ等しいと考えます。このことを**大数の法則**といいます。

共通テストでは，問題文に必ず $p \fallingdotseq R$ とするタイミングが書かれているので，誘導に従いましょう。

標本が十分大きく，母集団は標本に対して十分大きいことが前提ですが，標本調査の場合それらは満たしています。

標本において，ある事象 $A$ が起こるか起こらないかは独立と考えることができます。

その標本から，1 つ 1 つのデータを取り出すことを考えたとき，取り出したデータはある事象 $A$ が「起きる」か「起きない」かの 2 通りですね。

ってことは，抽出したデータの中からある事象 $A$ を満たす**個数 $X$ は二項分布に従う**と考えられます。

二項分布に従うことがわかれば，平均，分散，標準偏差はあっという間に求まりますね。

標本の大きさ $n$，標本比率 $R=p$（母比率と仮定）のとき，ある事象を満たしている個数 $X$ の平均 $E(X)$ は

$$E(X)=np$$

二項分布の平均と分散です。

となり，分散 $V(X)$ は

$$V(X)=np(1-p)$$

となります。

また，母集団から抽出した標本の大きさ $n$ は十分に大きいので，ある事象が起こる回数 $X$ は正規分布にも従います。

さらに，標本比率 $R$ は，$R=\dfrac{X}{n}$ で求めることができるので，$R$ の平均 $E(R)$ は

$$\begin{aligned}E(R)&=E\left(\frac{X}{n}\right)\\&=\frac{1}{n}E(X)\\&=\frac{1}{n}\cdot np=p\end{aligned}$$

となり，分散 $V(R)$ は

$$\begin{aligned}V(R)&=V\left(\frac{X}{n}\right)\\&=\frac{1}{n^2}V(X)\\&=\frac{1}{n^2}\cdot np(1-p)=\frac{p(1-p)}{n}\end{aligned}$$

と求められます。$X$ が正規分布に従うので，$R=\dfrac{X}{n}$ とおいた $R$ も正規分布に従います。

**母比率が従う分布**

標本比率 $R$ は，平均が $p$，分散が $\dfrac{p(1-p)}{n}$ の正規分布 $N\left(p,\ \dfrac{p(1-p)}{n}\right)$ に近似的に従います！

## 4 やっぱり具体例で理解しようよ！

これも具体例を通じて求める手順を理解してもらいます。

**例題**

ある県の小学生 400 名を無作為に選んだところ，眼鏡をかけている者は 80 名であった。この県の小学生について，眼鏡をかけている者の割合（母比率）$p$ を推定しよう。

標本比率は $R=\dfrac{80}{400}=0.2$ です。

これをもとに，信頼度 95 %の信頼区間を求めてみます。

① $R$ は正規分布 $N\left(p,\ \dfrac{p(1-p)}{400}\right)$ に近似的に従い，平均 $E(R)$ と標準偏差 $\sigma(R)$ は

$$E(R)=p\,,\qquad \sigma(R)=\sqrt{\frac{p(1-p)}{n}}=\sqrt{\frac{p(1-p)}{400}}$$

ただし，$p$ は $R=0.2$ にほぼ等しいので

$$\sigma(R)=\sqrt{\frac{0.2(1-0.2)}{400}}=\frac{\sqrt{0.16}}{20}=0.02$$

としてOKです。

② $z=\dfrac{R-E(R)}{\sigma(R)}=\dfrac{R-p}{0.02}$ とおくと，$z$ は標準正規分布に従うので，正規分布表が活用できます。

③ $P(-z_0\leqq z\leqq z_0)=0.95$ となる $z_0$ を正規分布表から読み取ると，$z_0=1.96$ となります。

④ 区間 $-1.96\leqq z\leqq 1.96$ を，$z=\dfrac{0.2-p}{0.02}$ を用いて書き直します。

$$-1.96\leqq\frac{0.2-p}{0.02}\leqq 1.96$$

整理して小数第 4 位を四捨五入すると

$$0.161\leqq p\leqq 0.239$$

となります。

以上のことから，小学生のうち眼鏡をかけている者の割合は，確率 95 %でこの範囲に入っているってことがわかりました。

めでたし，めでたし。