統計的な推測【大数の法則】p.98

青チャートB: 例題 88

【態度目標】しゃべる、質問する、説明する、動く、協力する、貢献する

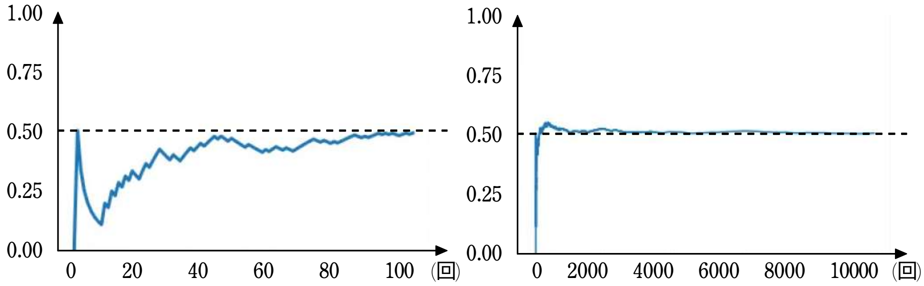
【内容目標】大数の法則の考え方を理解しよう

□大数（たいすう）の法則

例えば表が出る確率が $\frac{1}{2}$ の硬貨を何回も投げた結果を

確率としてグラフにしたものが下のようになっている。

数学的確率：計算で求める
統計的（経験的）確率：過去の統計から判断する

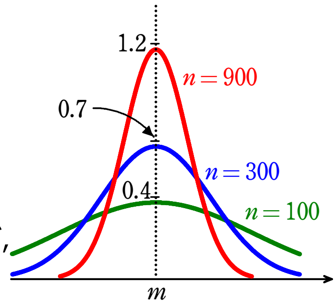


このように，試行の回数を増やしていくと統計的確率（経験的確率・統計学的確率）が数学的確率（公理的確率）に近づいていく（矛盾しない）。

「標本平均の分布と正規分布」について，$\sigma=10$ として，

$n=$ 100，300，900

の各場合に対して，近似される正規分布 $N\left(m,\ \frac{\sigma^2}{n}\right)$ を図示すると，右の図のようになる。この図から，n が大きくなればなるほど，$\overline{X}$ の分布は母平均 m の近くに集中して分布することがわかる。



一般に，標本の大きさ n を限りなく大きくしていくと，標本平均 $\overline{X}$ の分布を近似する正規分布の標準偏差 $\frac{\sigma}{\sqrt{n}}$ は，限りなく 0 に近づき，$\overline{X}$ は母平均 m の近くに限りなく集中して分布するようになる。したがって，$\overline{X}$ が m に近い値をとる確率が 1 に近づく。

大数の法則

すなわち，次の **大数の法則** が成り立つ。

大数の法則

母平均 m の母集団から大きさ n の無作為標本を抽出するとき，n が大きくなるに従って，その標本平均 $\overline{X}$ はほとんど確実に母平均 m に近づく。

標本数が大きくなれば，
標本平均は母平均に近づいていくということ

数学的確率と統計的（経験的）確率が
一致していく

統計的な推測【大数の法則】 p.98

青チャートB: 例題88

○サンプルの平均値と元の集団の平均値が近似できることを示しているものが**大数の法則**

○サンプルの平均値と元の集団の平均値の差が近づいていく過程をそのばらつき（分散）とともに示しているものが**中心極限定理**

と考えることができる。大数の法則をさらに深堀りしたものが中心極限定理というように理解するとイメージがつきやすいと思われる。

大数の法則から
中心極限定理へ

【発展】

厳密には大数の法則は収束をどのようにとらえるかに応じて

■ヤコブ・ベルヌーイによる大数の弱法則（WLLN：Weak Law of Large Numbers）

■エミール・ボレルやアンドレイ・コルモゴロフによる

大数の強法則（SLLN：Strong Law of Large Numbers）

の２つに大別される。どちらを指しているかは文脈により判断する必要がある。

確率変数 X_1，X_2，……，X_n が互いに独立で，平均値 m，標準偏差 σ の同じ確率分布に従うとする。このとき，X_1，X_2，……，X_n の標本平均 $\overline{X}$ は，任意の正数 ε に対して，

■大数の弱法則（WLLN：Weak Law of Large Numbers）

$$\lim_{n\to\infty} P(|\overline{X}-m| \geqq \varepsilon)=0$$

標本数 n が十分に大きくなれば，「独立に同一の分布に従う確率変数の標本平均と母平均のずれ」が一定の幅を超える確率が０に収束するという法則

（数学的には，標本平均は母平均に確率収束するという）

■大数の強法則（SLLN：Strong Law of Large Numbers）

$$\lim_{n\to\infty} P(|\overline{X}-m| < \varepsilon)=1$$

または　$P\left(\lim_{n\to\infty} \overline{X}=m\right)=1$

標本数 n が十分に大きくなれば，「独立に同一の分布に従う確率変数の標本平均」が母平均が一致する確率が１に近づくという法則

（数学的には，標本平均は母平均に概収束するという）

どちらにしても

「標本数が大きくなれば，標本平均は母平均に近づいていく（ m からの離れ方がだんだん小さくなっていく）」

ということをいっている

統計的な推測【大数の法則】 p.98

青チャートB: 例題 88

問８） 硬貨を n 回投げるとき，表の出る相対度数を R とする。

$n=100$ の場合について，$P\left(\left|R-\frac{1}{2}\right|\leqq 0.05\right)$ の値を，正規分布表を用いて求めよ。

解答　相対度数 R は，標本比率と同じ分布に従うから，

R は近似的に正規分布 $N\left(\frac{1}{2},\ \frac{1}{4n}\right)$ に従う。

標本比率 R は n が大きいとき近似的に正規分布 $N\left(p,\ \frac{pq}{n}\right)$ に従うとみなすことができる

$m=\frac{1}{2}$，$\sigma=\frac{1}{2\sqrt{n}}$ で標準化

よって，$Z=\dfrac{R-\frac{1}{2}}{\frac{1}{2\sqrt{n}}}$ は近似的に $N(0,\ 1)$ に従うから

$$P\left(\left|R-\frac{1}{2}\right|\leqq 0.05\right)=P\left(\left|\frac{Z}{2\sqrt{n}}\right|\leqq 0.05\right)$$
$$=P\left(-0.05\leqq\frac{Z}{2\sqrt{n}}\leqq 0.05\right)$$
$$=P(-0.05\times 2\sqrt{n}\leqq Z\leqq 0.05\times 2\sqrt{n})$$
$$=P(-0.1\sqrt{n}\leqq Z\leqq 0.1\sqrt{n})$$

ゆえに，$n=100$ のとき

$$P(-1\leqq Z\leqq 1)=2P(0\leqq Z\leqq 1)=2p(1)=0.6826$$

補足　硬貨を投げるという反復試行を n 回行うとき，表の出る回数を X とすると，

相対度数は $R=\frac{X}{n}$ で表される。

このとき，ε は任意の正数であるから，どんなに小さい正数 ε に対しても，

n が大きくなるにつれて期待値 $m=\frac{1}{2}$ に近づいていくわけであるから

n が無限に大きくなると「R と m との差が小さくなっていく確率」が 1 に近づいていく，

つまり $P\left(\left|R-\frac{1}{2}\right|\leqq\varepsilon\right)\to 1$ となることを示している。