13.1 見出しを読み解く

ワークシート

問題

STEP I 1992、問 14

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

step.maths.org/questions/92-s1-q14

ヒント

まずは自力で問題に取り組んでみよう。行き詰まったら、ヒントを**一つずつ**読み、そのたびに問題に戻ること。**手を止めて、やってみよう**と記したヒントが重要な段階である。各ヒントの後には計算用の余白がある。

? **考えてみよう**　死亡が一定の割合で、互いに独立かつランダムに起こるとすれば、年間の総数はどの程度ばらつくはずだろうか。平均が年 1080 人であるとき、年ごとの典型的なずれはどれくらいか。

段階 1　「飛行機は歩くより 6 倍安全」

扱う内容：小問 (i)

? **考えてみよう**　比 $1080/180 = 6$ は死亡者の総数を比べたものである。それはリスク（risk）を比べることと同じだろうか。公平なリスクの尺度を作るには、何を考慮に入れる必要があるだろうか。

段階 2 「空はより安全に：1990 年の死者はわずか 125 人」

扱う内容：小問 (ii)

考えてみよう 航空事故による死亡は、互いに独立（independent）なまれな事象だろうか。それとも、まとまって起こるものだろうか。年間の合計はどの程度ばらつくと考えられるか。1 年分のデータで傾向（trend）を示すことができるだろうか。

段階 3 「道路の惨状は悪化の一途：火曜日に 7 人死亡」

扱う内容：小問 (iii)

手を止めて、やってみよう 歩行者の死者は 1 日あたり平均何人か。1 日の死者数がその平均をもつポアソン分布（Poisson distribution）に従うとすると、死者が 7 人以上となる日は 1 年に何日あると期待されるか。また、1 日分のデータで何かが増加していると示すことができるだろうか。

段階 4 「憂慮すべき増加：1990 年に歩行者 1350 人死亡」

扱う内容：小問 (iv)

> **考えてみよう** 平均 1080 のポアソンモデルのもとで、1350 は平均より標準偏差（standard deviation）のいくつ分だけ大きいか。偶然でないとすれば、ほかにどのような説明が考えられるか。16 年間の平均が何を覆い隠しうるかを考えよ。

背後にある数学

シグナルとノイズ

統計学の問いはすべて、突き詰めれば**シグナル**（signal、本当の効果や傾向）を**ノイズ**（noise、ランダムな変動）から切り分ける問題である。ポアソンモデルは、互いに独立な事象の回数に含まれるノイズを測る物差しを与える。すなわち標準偏差 $\sqrt{\mu}$ である。稀な事象は**相対的な**変動が大きい。9 という回数はおよそ ± 3、つまりその大きさの 3 分の 1 ほど変動するが、10 000 という回数の変動はわずか 1% ほどである。小さな数（1 日の死者数、1 年の墜落回数）がしばしば過大に解釈されるのも、小さな町でのがんの「集積」がたいてい偶然だと判明するのも、このためである。

平均への回帰

さらにもう 1 つ落とし穴がある。交通安全カメラは、直前の 1 年間に異常に事故が多かった場所に設置されることが多い。翌年に事故が減ると、カメラの手柄とされる。しかし異常に悪い年の後には、いずれにせよもっと普通の年が続く可能性が高い。これが**平均への回帰**（regression to the mean）であり、1880 年代にゴルトン（Galton）が初めて記述したものである。カメラの効果を平均への回帰から切り分けるには、カメラが設置されなかった同様の場所と比較する必要がある。これが対照実験（controlled experiment）の背後にある考え方である。

ナイチンゲールの統計

クリミア戦争での看護師として名高いナイチンゲール（Florence Nightingale）は、先駆的な統計学者でもあった。1858 年の彼女の「鶏頭図（rose diagram）」は、負傷よりも予防可能な病気で亡くなる兵士のほうがはるかに多いことを示し、政府を説得して陸軍の衛生制度を改革させた。彼女は生の総数ではなく割合（兵士 1000 人あたりの年間死者数）を用いること、そして同じ条件のものどうしを比べることを強く主張した。これはまさにこの問題が問うている点である。彼女は王立統計学会（Royal Statistical Society）のフェローに選ばれた最初の女性であった。

次の一歩

1. **ポアソン分布の裾。** X が平均 2.96 のポアソン分布に従うとき、$\mathrm{P}(X \geq 7) \approx 0.031$ であることを確かめよ。歩行者の死者が 9 人以上となる日は、1 年に何日あると期待されるか。

2. **まとまって起こる死亡。** 航空機の墜落が年 1.2 回のポアソン過程（Poisson process）として起こり、1 回の墜落ごとに 150 人が亡くなるとする。航空機事故による年間死者数の平均と標準偏差を求めよ。死者が 125 人以下という年は珍しいか。死者がちょうど 125 人という年はありうるか。

3. **有意な増加か。** 1974–1989 年の数値が 1080 ± 33 で一定であったとすると、ある 1 年に 1350 人以上となる確率を（正規近似（normal approximation）を用いて）求めよ。

4. **よりよい見出しを書く。** 4 つの見出しそれぞれについて、データが実際に裏づける見出しを書け。

さらに学ぶために

- **David Spiegelhalter, *The Art of Statistics***（Pelican）。データから学ぶ方法を、ニュースからの多くの例とともに解説する。
- **Tim Harford, *How to Make the World Add Up***（Bridge Street）。メディアに現れる数字について考えるための 10 のルール。
- **Darrell Huff, *How to Lie with Statistics***（Penguin）。1954 年に書かれた短い古典で、今なお色あせない。

詳しい解説付きの解答は第 13.1 節にある：

https://beyondstep.keithmcnulty.org/ja/ch13/headlines.html

毎週 1 問の STEP の問題を丁寧に解き明かす The STEP Up。メールで無料配信中：keithmcnulty.substack.com