解答例

STEP I 1992、第 14 問

確率・統計 ・ 統計的主張の解釈

> **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

この問題は、*Beyond STEP* の第 13.1 節「**見出しを読み解く**」で詳しく解説されている：
https://beyondstep.keithmcnulty.org/ja/ch13/headlines.html

鍵となる考え方

どの見出しも、ある一つの数値から一足飛びに結論を導いている。それぞれについて、二つの問いを立てるとよい。第一に、同じ種類のものどうしを比べているか。すなわち、同じものを同じ方法で測り、リスクへの曝露量（exposure）に応じて調整しているか。第二に、その差は単なる偶然で説明できないか。第二の問いについては、まれで互いに独立な事象の個数に対するポアソンモデル（Poisson model）が自然な物差しとなる。その標準偏差（standard deviation）は平均の平方根である。このモデルがいつ当てはまり、いつ当てはまらないかを見極めることが、(ii) から (iv) の核心である。

解答

以下では一貫して、「平均して」とは 1974–1989 年の年平均を意味する。すなわち、歩行者の死者は年 1080 人、航空機による死者は年 180 人である。

小問 (i)

「6 倍」という数字は比 $1080/180 = 6$ に由来するが、これはリスクではなく単なる総数どうしを比べたものである。リスクとは曝露量の単位あたりの死者数であり、1 回の移動あたり、1 時間あたり、あるいは移動距離 1 キロメートルあたりなどで測る。ポルダビアではほぼすべての人が毎日道路の近くを歩くが、飛行機に乗る人は比較的少な

く、乗るとしてもたまにである。安全性を比べるには、次のようなものが必要である。

$$\frac{\text{歩行者の死者数}}{\text{歩いた総時間（または総距離 km）}} \quad \text{対} \quad \frac{\text{航空機による死者数}}{\text{飛行した総時間（または総距離 km）}},$$

しかし見出しからは、どちらの分母も得られない。測り方によっては、結論が逆になることさえありうる。距離あたりで見れば飛行機は非常に安全に見えるが、時間あたりや 1 回の移動あたりで見れば、それほど安全には見えないかもしれない。

ほかにも問題がある。ここでいう「歩くこと」の実態は、徒歩中に車両にはねられることであり、この数値は人通りの少ない場所の歩行者については何も語らない。航空機による死者には乗客だけでなく乗務員も含まれる。また、ある期間の平均値によって 1991 年の安全性について何かを「証明」することはできない。せいぜい 1974–1989 年についての証拠になるだけである。

小問 (ii)

一見すると、125 は平均の 180 を大きく下回っている。この減少は偶然以上のものだろうか。

もし死亡が一人ずつ独立に起こるのであれば、平均 180 のポアソンモデルが妥当であろう。その標準偏差は $\sqrt{180} \approx 13.4$ であるから、125 は平均より約 4.1 標準偏差下にあり、偶然に起こる可能性は非常に低い（$P(X \leq 125) \approx 10^{-5}$）。

しかし、航空機による死亡は一人ずつ起こるのではない。死亡はまとまって生じ、一度の墜落で数十人から数百人が亡くなりうる。自然なモデルは、**死亡事故**の件数がポアソン分布に従い、各事故がランダムな人数の死者を出すというものである。これにより、ばらつきは大幅に大きくなる。例として、死亡事故が平均して年 6 件あり、各事故で 30 人が亡くなるとしよう。すると年間の死者数は $N \sim \text{Po}(6)$ として $30N$ となり、その標準偏差は $30\sqrt{6} \approx 73$ であって、

$$P(\text{死者数} \leq 125) = P(N \leq 4) \approx 0.29.$$

この（説明のための）数値のもとでは、死者が 125 人の年はまったく普通である。事故が通常より 1 件か 2 件少なければよいだけだからである。各事故の規模もばらつくとすれば、散らばりはさらに大きくなる。

したがって、ある一年の数値からわかることはごくわずかである。1990 年にどれだけ飛行が行われたか（航空交通量が減っていれば、いずれにせよ死者は減ると予想される）、さらに真の傾向があるかどうかを見るために複数年の数値も知りたいところである。見出しの主張は裏づけられていない。

小問 (iii)

ここには、別々の誤りが二つある。

一日だけでは増加を示せない。「増えている」とは時間的な傾向についての主張であり、ある一日の数値はまったく比較を与えない。長期間にわたる日ごと（あるいは年ごと）の数値が必要である。

7人はそもそも珍しくない。手元にある道路関係の数値は歩行者についてのものだけである。すなわち年 1080 人、つまり

$$\lambda = \frac{1080}{365} \approx 2.96 \text{ per day.}$$

歩行者の死亡はポアソン分布でかなりよくモデル化できる。まれであり、ほとんどが一つの事故につき一人であり、おおむね独立だからである。ある日に亡くなる歩行者の数を $X \sim \mathrm{Po}(2.96)$ とする。すると

$$P(X \geq 7) = 1 - \sum_{k=0}^{6} \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \approx 0.031,$$

となるから、歩行者が少なくとも 7 人亡くなる日は、年におよそ $365 \times 0.031 \approx 11$ 日あると予想される。しかも見出しは「死亡した人」と言っており、運転者、同乗者、自転車に乗る人も含まれる。道路での一日あたりの総死者数の率は 2.96 より大きいはずであり、7 人という数はいっそう目立たないものになる。

> **考えてみよう** 平均の 2 倍を超える死者が出た日は、憂慮すべきではないのか。

小さな個数については、そうではない。平均が約 3 のとき、ポアソン分布の標準偏差は $\sqrt{3} \approx 1.7$ であり、平均に比べて大きいので、日によって比率として大きく振れるのは普通のことである。最悪の日を選び出すことは選択効果（selection effect）でもある。一年には 365 日あり、新聞はいつでもひどい日を見つけられる。曜日も関係する。交通量や徒歩で外出する人の数は、曜日や季節によって変わるからである。一定の率をもつポアソンモデルは、大まかな第一近似にすぎない。

小問 (iv)

これは、本当に検討に値する見出しである。まず統計を、次に解釈を考える。

1350 は偶然以上のものか。ここでは (iii) で述べた理由により、ポアソンモデルがよい出発点となる。平均が 1080 のとき標準偏差は $\sqrt{1080} \approx 32.9$ であり、

$$\frac{1350 - 1080}{32.9} \approx 8.2.$$

平均より 8 標準偏差上の値は、このモデルのもとで 10^{-15} 程度の確率しかもたない。たとえ年ごとの真の散らばりがポアソン分布の値の 2 倍や 3 倍であったとしても（たとえば厳しい冬のために）、一定の率 1080 と比べれば 1350 はやはり非常に珍しい。したがって、何かが変化した可能性が非常に高い。問題は、それが何かである。

何と比べているのか。見出しは 1990 年を 1989 年とではなく、16 年間の平均と比べている。もし歩行者の死者数がこの期間に着実に増え続けていたなら、平均 1080 は期間の中ごろ（1981–82 年ごろ）を表すことになり、1990 年はそれを大きく上回るはずである。たとえば、1974 年の約 855 人から毎年 30 人ずつ着実に増えると、平均は 1080 となり、1990 年には約 1335 人となる。この場合、1990 年は「急増」などではまったくなく、長期的な傾向が続いているだけである。その傾向は見出しに値するかもしれないが、それは別の見出しである。二つの筋書きを見分けるには、年ごとの数値が必要である。

曝露量は変化したか。人口が増えたり、車両が増えたり、歩く量が増えたりすれば、1 回の移動あたりの歩行の危険性が少しも増えなくても、死者数は増える。公正な比較は、人口 100 000 人あたりや歩行距離 1 キロメートルあたりの死者数のような率によるものである。

測り方は変化したか。死亡の記録や分類の方法が変われば、見かけ上の急増が生じる。たとえば、事故後どれだけの期間内の死亡を交通事故死として数えるか、誰を歩行者として数えるかが変わった場合や、報告がより正確になった場合である。そのような変化がないことを確かめる必要がある。

言葉の選び方は正しいか。「死傷者（casualties）」とは通常、死亡した人**または**負傷した人を意味する。引用されている数値は死者だけのものであるから、見出しはこの数値が示す以上のことを言っている。負傷者ははるかに多く、まったく異なる傾向を示すかもしれない。

まとめると、1350 への増加は、一定の率 1080 のまわりのポアソン分布による変動で説明するにはあまりにも大きい。しかし、それを「憂慮すべき」と呼ぶ前に、最近の年ごとの数値（突然の増加と長期的傾向を区別するため）と、人口や交通量の数値（個数を率に直すため）が必要である。さらに、定義が変わっていないことも確かめる必要がある。

検算

確率は、ポアソン級数を正確に足し合わせる方法と、2×10^5 回の試行によるシミュレーションとで確かめた。(iii) では、$\lambda = 1080/365$ のとき $P(X \geq 7) = 0.0315$ である（シミュレーションでは 0.0317）。(ii) では、各 30 人の死者を出す事故が年 6 件の場合、$P(N \leq 4) = 0.285$ である。事故の規模が 1 から 59 まで一様にばらつくとすると、シミュレーションにより標準偏差は 84、$P(\text{死者数} \leq 125) \approx 0.28$ となった。ポアソン分布に従う個数についての手早い目安は、平均からおよそ平方根 2 つ分以内の値は目立たないというものである。一日については $3 + 2 \times 1.7 \approx 6.4$ であるから、7 はその範囲をわずかに外れるだけであり、選べる日が 365 日あれば、そのような日は必ず起こる。

一年については $1080 + 2 \times 33 = 1146$ であり、1350 には遠く及ばない。

よくある間違い

- (i) で率ではなく総数を比べること。リスクには必ず分母が必要である。
- (ii) でポアソン分布の標準偏差 $\sqrt{180}$ を用いること。航空機事故による死亡はまとまって起こるので、散らばりははるかに大きい。
- (iv) で「8 標準偏差だから本当の増加だ」で止めてしまい、16 年間の平均との比較、人口、あるいは定義によって説明できないかを問わないこと。

STEP のその先へ

小問 (ii) は**過分散**（overdispersion）に関するものである。ポアソン分布に従う個数のまとまりからなり、各まとまりの大きさがランダムであるような総和は**複合ポアソン**（compound Poisson）確率変数であり、その分散は λ ではなく $\lambda\,\mathrm{E}[S^2]$ である。小問 (iv) は仮説検定（hypothesis testing）へ、さらに統計的に有意（statistically significant）な変化とは実際に何を意味するのかという、はるかに難しい問いへとつながる。どちらも大学の確率・統計の講義の早い段階で登場する。Blitzstein and Hwang, *Introduction to Probability*（確率論入門）は、ポアソンモデルとその限界について読むのによい本である。

Beyond STEP 解答集：https://beyondstep.keithmcnulty.org/ja/。解答本文は CC0 1.0 である。問題そのものは © Cambridge University Press & Assessment に帰属し、ここには掲載していない。英語版から機械翻訳したものであり、まだ校閲を経ていない。