解答例

STEP III 1996、第 13 問

確率・統計 ・ チェビシェフの不等式

> **STEP III 1996、問 13**
>
> 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/96-s3-q13

鍵となる考え方

分散（variance）は、各値に対応する非負の項 $p_i(x_i-\mu)^2$ の和である。この和を、平均（mean）から遠い値（平均から λ 以上離れているもの）と近い値とに分ける。遠い値をとる確率の**上界**（upper bound）を得るには、近い項を捨て、遠い各項を下から評価する。**下界**（lower bound）を得るにはその逆を行う。すなわち、すべての項を上から評価し、近い項は λ^2 で、遠い項は k^2 で抑える。

解答

定義と $aX+b$

$p_i=\mathrm{P}(X=x_i)$ とおくと、$p_i\ge 0$ かつ $\sum_{i=1}^{n} p_i=1$ である。このとき

$$\mathrm{E}(X)=\sum_{i=1}^{n} x_i p_i, \qquad \mathrm{var}(X)=\mathrm{E}((X-\mu)^2)=\sum_{i=1}^{n}(x_i-\mu)^2 p_i,$$

である。ただし $\mu=\mathrm{E}(X)$ とする。（同値な形として $\mathrm{var}(X)=\mathrm{E}(X^2)-\mu^2$ もあるが、ここでは最初の形を用いる。）

確率変数（random variable）$aX+b$ は確率 p_i で値 ax_i+b をとる。（$a=0$ のときはこれらの値が一致するが、等しい値の確率をまとめても以下の和は変わらない。）よって

$$\mathrm{E}(aX+b)=\sum_i (ax_i+b)p_i = a\sum_i x_i p_i + b\sum_i p_i = a\mathrm{E}(X)+b,$$

となる。ここで $\sum_i p_i=1$ を用いた。したがって $aX+b-\mathrm{E}(aX+b)=a(X-\mu)$ であり、

$$\mathrm{var}(aX+b)=\sum_i \big(a(x_i-\mu)\big)^2 p_i = a^2\sum_i (x_i-\mu)^2 p_i = a^2\,\mathrm{var}(X).$$

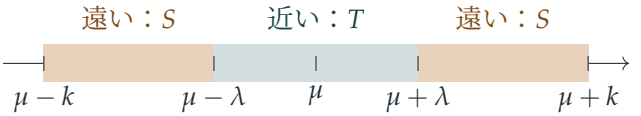
b を加えると、すべての値と平均がそろってずれるので、散らばりには影響しない。

上界

$\mu = \mathrm{E}(X)$ とし、$|x_i - \mu| \geq \lambda$ となる添字 i の集合（「遠い」値）を S、残り、すなわち $|x_i - \mu| < \lambda$ となるものを T とする。このとき

$$\mathrm{P}(|X - \mu| \geq \lambda) = \sum_{i \in S} p_i.$$

この確率を P とおく。



分散の各項は非負であるから、T の項を取り除いても和が大きくなることはない。また $i \in S$ に対しては $(x_i - \mu)^2 \geq \lambda^2$ である。したがって

$$\begin{aligned}\mathrm{var}(X) &= \sum_{i \in S} (x_i - \mu)^2 p_i + \sum_{i \in T} (x_i - \mu)^2 p_i \\ &\geq \sum_{i \in S} (x_i - \mu)^2 p_i \\ &\geq \lambda^2 \sum_{i \in S} p_i = \lambda^2 P.\end{aligned}$$

$\lambda > 0$ であるから λ^2 で割ることができ、

$$\mathrm{P}(|X - \mu| \geq \lambda) \leq \frac{\mathrm{var}(X)}{\lambda^2},$$

が得られる。これが示すべき不等式である。（S が空集合ならば S 上の和はどちらもゼロとなり、不等式は自明に成り立つ。）

下界

次に、さらにすべての i について $|x_i - \mu| \leq k$ が成り立つと仮定する。今度は分散のすべての項を**上から**評価する。

- $i \in S$ に対しては、新しい仮定より $(x_i - \mu)^2 \leq k^2$ である。
- $i \in T$ に対しては、T の定義より $(x_i - \mu)^2 < \lambda^2$ である。

$\sum_{i \in T} p_i = 1 - P$ であるから、次が得られる。

$$\mathrm{var}(X) \leq k^2 \sum_{i \in S} p_i + \lambda^2 \sum_{i \in T} p_i = k^2 P + \lambda^2 (1 - P) = \lambda^2 + (k^2 - \lambda^2) P.$$

整理すると

$$(k^2-\lambda^2)\,P \geq \operatorname{var}(X)-\lambda^2.$$

考えてみよう $k^2-\lambda^2$ で割ってよいだろうか？

条件より $k \geq \lambda > 0$ であるから $k^2-\lambda^2 \geq 0$ である。$k > \lambda$ ならばこれは正であり、割っても不等号の向きは変わらない。

$$\mathrm{P}\big(|X-\mu| \geq \lambda\big) \geq \frac{\operatorname{var}(X)-\lambda^2}{k^2-\lambda^2},$$

これが求める結果である。境界の場合 $k=\lambda$ では右辺の分母がゼロとなり、この主張は意味をもたない。その直前の不等式は依然として成り立ち、$\operatorname{var}(X) \leq \lambda^2$ を意味するが、これは各 $(x_i-\mu)^2 \leq k^2 = \lambda^2$ であることから明らかである。

$\operatorname{var}(X) \leq \lambda^2$ のときは下界はゼロ以下となり、新しいことは何も述べていない。この下界が意味をもつのは分散が λ^2 に比べて大きい場合だけであり、そのとき確率の一部が分布の裾（tails）へ押し出されることになる。

検算

- $X = \pm 1$ となる確率がそれぞれ 0.4、± 4 となる確率がそれぞれ 0.1 であるとする。すると $\mu = 0$、$\operatorname{var}(X) = 0.8\times 1 + 0.2\times 16 = 4$ である。$\lambda = 1$、$k = 4$ とすると $P = 1$ であり、上界は 4、下界は $\frac{3}{15} = \frac{1}{5}$ となって、確かに $\frac{1}{5} \leq 1 \leq 4$ が成り立つ。
- $X = 0$ または $\pm c$ であるとし、それぞれ $\mathrm{P}(X = \pm c) = q/2$ とする。すると $\operatorname{var}(X) = qc^2$ であり、$\lambda = c$ とすると上界はちょうど $q = P$ に一致する。つまりチェビシェフ（Chebyshev）の不等式（Chebyshev's inequality）は一般にはこれ以上改良できない。
- 両方の評価式、および $\mathrm{E}(aX+b)$ と $\operatorname{var}(aX+b)$ の公式を、とる値が高々七つである 20 000 個のランダムな分布について厳密な有理数演算で検証したところ、成り立たない例は一つもなかった。

よくある間違い

- $\operatorname{var}(aX+b) = a\operatorname{var}(X)$ や $a^2\operatorname{var}(X)+b$ と書いてしまう。分散は偏差について二次であり、平行移動はどの偏差も変えない。
- 下界の評価で、近い項を λ^2 ではなく 0 で抑えてしまう。すると $\operatorname{var}(X) \leq k^2 P$ が得られるが、これは正しいものの、求められている結果ではない。

- $k^2 - \lambda^2$ で割る際に、それが正であること（および $k = \lambda$ を除外しなければならないこと）に言及しない。

STEP のその先へ

最初の不等式はチェビシェフの不等式であり、項を捨てるという技法は、$Y \geq 0$ に対するマルコフ（Markov）の不等式（Markov's inequality）$\mathrm{P}(Y \geq c) \leq \mathrm{E}(Y)/c$ の背後にあるものと同じである（これを $Y = (X - \mu)^2$ に適用すればよい）。大学 1 年の確率論の講義では、チェビシェフの不等式を使うと大数の弱法則（weak law of large numbers）がわずか 2 行で証明できる。すなわち、X の独立なコピー n 個の平均は分散 $\mathrm{var}(X)/n$ をもつので、μ から大きく外れる可能性は低い。2 つ目の不等式は有界性を必要とする「逆向き」の版であり、ペイリー＝ジグムントの不等式（Paley–Zygmund inequality）と同じ精神に立つものである。Grimmett and Welsh の *Probability: An Introduction*（確率論入門）がこれらすべてを扱っている。

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。