詳細解答

STEP III 1996，第 13 題

概率與統計 • 切比雪夫不等式

> **STEP III 1996，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/96-s3-q13

關鍵思路

方差（variance）是若干個非負項 $p_i(x_i-\mu)^2$ 之和，每個取值對應一項。把這個和分成兩部分：遠離平均值（mean）的取值（相距至少 λ）和靠近平均值的取值。要得到「遠」的概率的上界（upper bound），就把「近」的各項丟掉，並把每個「遠」的項從下方估計。要得到下界（lower bound），則反其道而行：把每一項都從上方估計，「近」的項不超過 λ^2，「遠」的項不超過 k^2。

解答

定義與 $aX+b$

記 $p_i=\mathrm{P}(X=x_i)$，於是 $p_i\geq 0$ 且 $\sum_{i=1}^n p_i=1$。那麼

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

其中 $\mu=\mathrm{E}(X)$。（等價地，$\mathrm{var}(X)=\mathrm{E}(X^2)-\mu^2$，但我們將會使用的是第一種形式。）

隨機變量（random variable）$aX+b$ 以概率 p_i 取值 ax_i+b。（若 $a=0$，這些值會重合，但把相等取值的概率相加並不會改變下面的各個和。）所以

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

這裏用到 $\sum_i p_i=1$。因此 $aX+b-\mathrm{E}(aX+b)=a(X-\mu)$，而

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

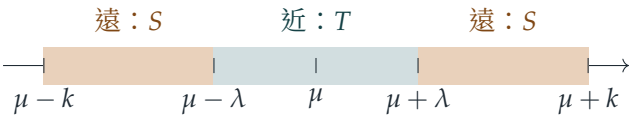
加上 b 會把每個取值和平均值一同平移，所以對離散程度毫無影響。

上界

設 $\mu = \mathrm{E}(X)$，設 S 為滿足 $|x_i - \mu| \geq \lambda$ 的下標 i 所組成的集合（即「遠」的取值），並設 T 為其餘下標的集合，即 $|x_i - \mu| < \lambda$ 者。那麼

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

把這個概率記作 P。



方差的每一項都是非負的，所以丟掉 T 中的各項只會令和變小；而對於 $i \in S$，我們有 $(x_i - \mu)^2 \geq \lambda^2$。因此

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

由於 $\lambda > 0$，我們可以除以 λ^2，得到

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

正是所求。（若 S 是空集，則兩個對 S 求的和都是零，不等式顯然成立。）

下界

現在再假設對每個 i 都有 $|x_i - \mu| \leq k$。這一次我們把方差的每一項都從上方估計：

- 對於 $i \in S$，由新的假設可知 $(x_i - \mu)^2 \leq k^2$；
- 對於 $i \in T$，由 T 的定義可知 $(x_i - \mu)^2 < \lambda^2$。

由於 $\sum_{i \in T} p_i = 1 - P$，由此得

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

移項整理，

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

想一想 我們可以除以 $k^2-\lambda^2$ 嗎？

題目告訴我們 $k \geq \lambda > 0$，所以 $k^2-\lambda^2 \geq 0$。若 $k > \lambda$，它嚴格為正，相除後不等號方向不變：

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

這就是所要求的結果。在邊界情況 $k = \lambda$ 時，右邊的分母為零，這個命題沒有意義；但它前面的那個不等式仍然成立，並給出 $\mathrm{var}(X) \leq \lambda^2$，這是顯然的，因為每個 $(x_i-\mu)^2 \leq k^2 = \lambda^2$。

當 $\mathrm{var}(X) \leq \lambda^2$ 時，下界為零或負數，因而沒有帶來任何新資訊；只有當方差相對於 λ^2 較大，迫使部分概率落到兩端的尾部時，它才有實質內容。

驗算

- 取 $X = \pm 1$，概率各為 0.4，以及 ± 4，概率各為 0.1，於是 $\mu = 0$ 且 $\mathrm{var}(X) = 0.8 \times 1 + 0.2 \times 16 = 4$。取 $\lambda = 1, k = 4$：$P = 1$，上界是 4，下界是 $\frac{3}{15} = \frac{1}{5}$，而確實有 $\frac{1}{5} \leq 1 \leq 4$。
- 取 $X = 0$ 或 $\pm c$，其中 $\mathrm{P}(X = \pm c) = q/2$（兩者各自如此）。那麼 $\mathrm{var}(X) = qc^2$，而取 $\lambda = c$ 時，上界恰好是 $q = P$：一般而言，切比雪夫不等式（Chebyshev's inequality）無法再改進。
- 我們以精確的有理數運算，在 20 000 個取值最多七個的隨機分佈上，檢驗了兩個界以及 $\mathrm{E}(aX+b)$ 和 $\mathrm{var}(aX+b)$ 的公式；沒有一例出錯。

常見錯誤

- 寫成 $\mathrm{var}(aX+b) = a\,\mathrm{var}(X)$ 或 $a^2\mathrm{var}(X)+b$：方差是偏差的二次式，而平移不會改變任何偏差。
- 在下界中，把「近」的各項以 0 而非 λ^2 作上界。這樣會得到 $\mathrm{var}(X) \leq k^2 P$，雖然正確，卻不是題目所要求的。
- 除以 $k^2-\lambda^2$ 而沒有指出它是正數（以及必須排除 $k = \lambda$ 的情況）。

STEP 之外

第一個不等式就是切比雪夫不等式，而丟棄若干項的技巧，正是馬爾可夫不等式（Markov's inequality）$\mathrm{P}(Y \geq c) \leq \mathrm{E}(Y)/c$（適用於 $Y \geq 0$）背後的同一手法（把它應

用於 $Y = (X - \mu)^2$ 即可）。在大學一年級的概率課程中，切比雪夫不等式能用兩行證明弱大數定律（weak law of large numbers）：X 的 n 個獨立複本的平均值，其方差為 $\mathrm{var}(X)/n$，所以它不大可能遠離 μ。第二個不等式是一個需要有界性（boundedness）的「反向」版本，與 Paley–Zygmund 不等式（Paley–Zygmund inequality）一脈相承。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教材）涵蓋了以上全部內容。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。