详细解答

STEP III 1996，第 13 题

概率与统计 · 切比雪夫不等式

> **STEP III 1996，第 13 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/96-s3-q13

关键思路

方差（variance）是若干个非负项 $p_i(x_i-\mu)^2$ 之和，每个取值对应一项。把这个和分成两部分：远离平均值（mean）的取值（相距至少 λ）和靠近平均值的取值。要得到“远”的概率的上界（upper bound），就把“近”的各项丢掉，并把每个“远”的项从下方估计。要得到下界（lower bound），则反其道而行：把每一项都从上方估计，“近”的项不超过 λ^2，“远”的项不超过 k^2。

解答

定义与 $aX+b$

记 $p_i=\mathrm{P}(X=x_i)$，于是 $p_i\geq 0$ 且 $\sum_{i=1}^{n} p_i=1$。那么

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

其中 $\mu=\mathrm{E}(X)$。（等价地，$\mathrm{var}(X)=\mathrm{E}(X^2)-\mu^2$，但我们将会使用的是第一种形式。）

随机变量（random variable）$aX+b$ 以概率 p_i 取值 ax_i+b。（若 $a=0$，这些值会重合，但把相等取值的概率相加并不会改变下面的各个和。）所以

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

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

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

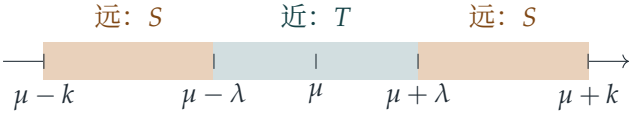
加上 b 会把每个取值和平均值一同平移，所以对离散程度毫无影响。

上界

设 $\mu = \mathrm{E}(X)$，设 S 为满足 $|x_i - \mu| \geq \lambda$ 的下标 i 所组成的集合（即“远”的取值），并设 T 为其余下标的集合，即 $|x_i - \mu| < \lambda$ 者。那么

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

把这个概率记作 P。



方差的每一项都是非负的，所以丢掉 T 中的各项只会令和变小；而对于 $i \in S$，我们有 $(x_i - \mu)^2 \geq \lambda^2$。因此

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

由于 $\lambda > 0$，我们可以除以 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，则两个对 S 求的和都是零，不等式显然成立。）

下界

现在再假设对每个 i 都有 $|x_i - \mu| \leq k$。这一次我们把方差的每一项都从上方估计：

- 对于 $i \in S$，由新的假设可知 $(x_i - \mu)^2 \leq k^2$；
- 对于 $i \in T$，由 T 的定义可知 $(x_i - \mu)^2 < \lambda^2$。

由于 $\sum_{i \in T} p_i = 1 - P$，由此得

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

移项整理，

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

想一想 我们可以除以 $k^2-\lambda^2$ 吗?

题目告诉我们 $k \geq \lambda > 0$，所以 $k^2-\lambda^2 \geq 0$。若 $k > \lambda$，它严格为正，相除后不等号方向不变：

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

这就是所要求的结果。在边界情况 $k=\lambda$ 时,右边的分母为零,这个命题没有意义;但它前面的那个不等式仍然成立,并给出 $\mathrm{var}(X) \leq \lambda^2$,这是显然的,因为每个 $(x_i-\mu)^2 \leq k^2 = \lambda^2$。

当 $\mathrm{var}(X) \leq \lambda^2$ 时，下界为零或负数，因而没有带来任何新信息；只有当方差相对于 λ^2 较大，迫使部分概率落到两端的尾部时，它才有实质内容。

验算

- 取 $X = \pm 1$，概率各为 0.4，以及 ± 4，概率各为 0.1，于是 $\mu = 0$ 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 $\frac{3}{15} = \frac{1}{5}$，而确实有 $\frac{1}{5} \leq 1 \leq 4$。
- 取 $X = 0$ 或 $\pm c$，其中 $\mathrm{P}(X = \pm c) = q/2$（两者各自如此）。那么 $\mathrm{var}(X) = qc^2$，而取 $\lambda = c$ 时，上界恰好是 $q = P$：一般而言，切比雪夫不等式（Chebyshev's inequality）无法再改进。
- 我们以精确的有理数运算，在 20 000 个取值最多七个的随机分布上，检验了两个界以及 $\mathrm{E}(aX+b)$ 和 $\mathrm{var}(aX+b)$ 的公式；没有一例出错。

常见错误

- 写成 $\mathrm{var}(aX+b) = a\,\mathrm{var}(X)$ 或 $a^2\mathrm{var}(X)+b$：方差是偏差的二次式，而平移不会改变任何偏差。
- 在下界中，把“近”的各项以 0 而非 λ^2 作上界。这样会得到 $\mathrm{var}(X) \leq k^2 P$，虽然正确，却不是题目所要求的。
- 除以 $k^2-\lambda^2$ 而没有指出它是正数（以及必须排除 $k=\lambda$ 的情况）。

STEP 之外

第一个不等式就是切比雪夫不等式,而丢弃若干项的技巧,正是马尔可夫不等式(Markov's inequality) $\mathrm{P}(Y \geq c) \leq \mathrm{E}(Y)/c$（适用于 $Y \geq 0$）背后的同一手法（把它应用于 $Y = (X-\mu)^2$ 即可）。在大学一年级的概率课程中，切比雪夫不等式能用两行证明弱大数定律（weak law of large numbers）：X 的 n 个独立复本的平均值，其方差为 $\mathrm{var}(X)/n$，所以

它不大可能远离 μ。第二个不等式是一个需要有界性（boundedness）的“反向”版本，与 Paley–Zygmund 不等式（Paley–Zygmund inequality）一脉相承。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（概率论入门教材）涵盖了以上全部内容。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。