Law of Large Numbers

Renhe W.

1 Definitions and Proofs

证明之前，回顾一些引理：

Lemma 1.1 (Lévy 极限定理 (Lévy's Convergence Theorem)). *如果一个随机变量序列 $\{X_n\}$ 独立同分布，且存在常数 $E[X_n] = \mu$ 和 $\mathrm{Var}[X_n] = \sigma^2 < \infty$ ，那么序列 $\{\bar{X}_n\}$ ，其中 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$，满足*

$$\lim_{n\to\infty} P\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x),$$

其中 $\Phi(x)$ 是均值为 0 和方差为 1 的正态分布的累积分布函数.

统计学中常用的几个大数定理的定义如下：

Theorem 1.1 (伯努利大数定理 (Bernoulli's Law of Large Numbers)). *对于独立同分布 (i.i.d.) 的随机变量序列 $\{X_n\}$ ，每个 X_n 取值于 $\{0, 1\}$ ，且 $P(X_n = 1) = p$ 和 $P(X_n = 0) = 1 - p$ ，当 n 趋近于无穷大时，样本均值 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ 收敛于 p ，即:*

$$\lim_{n\to\infty} P\left(|\bar{X}_n - p| < \epsilon\right) = 1.$$

Proof. 伯努利大数定理的证明可以基于切比雪夫不等式。考虑到 X_n 是伯努利随机变量，我们有 $E[X_n] = p$ 和 $\mathrm{Var}[X_n] = p(1-p)$. 样本均值 $\bar{X}_n$ 的期望和方差分别为:

$$E\left[\bar{X}_n\right] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = p,$$

$$\mathrm{Var}\left[\bar{X}_n\right] = \mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{p(1-p)}{n},$$

根据切比雪夫不等式，我们有:

$$P\left(\left|\bar{X}_n - p\right| > \epsilon\right) \leq \frac{\mathrm{Var}\left[\bar{X}_n\right]}{\epsilon^2} = \frac{p(1-p)}{n\epsilon^2},$$

当 n 趋于无穷大时，上式的右边趋于 0 ，因此

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - p\right| > \epsilon\right) = 0,$$

这就证明了样本均值 $\bar{X}_n$ 依概率收敛于 p. □

Theorem 1.2 (辛钦大数定理 (Khintchine's Law of Large Numbers)). 对于一组独立同分布的随机变量 $\{X_n\}$ ，如果它们的期望 $E[X_n]$ 存在，则样本均值 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ 当 n 趋近于无穷大时依概率收敛于 $E[X_n]$ ，即对于任意 $\epsilon > 0$:

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - E[X_n]\right| > \epsilon\right) = 0.$$

Proof. 辛钦大数定理的证明通常基于特征函数或者矩生成函数. 这里我们给出一个基于特征函数的简化证明.

由于 $\{X_n\}$ 是独立同分布的，并且 $E[X_n]$ 存在，我们可以定义 $Y_n = X_n - E[X_n]$ 使得 $E[Y_n] = 0$. 对于任意实数 t ，特征函数 $\phi_{Y_n}(t)$ 存在，并且我们有:

$$\phi_{\bar{Y}_n}(t) = \left[\phi_{Y_n}\left(\frac{t}{n}\right)\right]^n,$$

由于 $E[Y_n] = 0$ ，我们可以将 $\phi_{Y_n}(t)$ 在 $t = 0$ 处泰勒展开:

$$\phi_{Y_n}(t) = 1 + itE[Y_n] - \frac{t^2}{2}E\left[Y_n^2\right] + o\left(t^2\right),$$

这里 $o(t^2)$ 表示当 t 趋近于 0 时比 t^2 更快地趋近于 0 的项. 将上述展开带入 $\phi_{Y_n}(t)$ 的表达式得到:

$$\phi_{Y_n}(t) = \left[1 - \frac{t^2}{2n^2}E\left[Y_n^2\right] + o\left(\frac{t^2}{n^2}\right)\right]^n,$$

当 n 趋近于无穷大时，由于 $E[Y_n^2] < \infty$ ，我们有:

$$\lim_{n\to\infty} \phi_{Y_n}(t) = e^{-\frac{t^2}{2}E\left[Y_n^2\right]},$$

上述极限就是均值为 0 、方差为 $E[Y_n^2]$ 的正态分布的特征函数. 由 Lévy 极限定理，$\bar{Y}_n$ 依分布收敛于均值为 0 、方差为 $E[Y_n^2]$ 的正态分布. 由于 $\bar{X}_n - E[X_n] = \bar{Y}_n$ ，我们得到 $\bar{X}_n$ 依分布收敛于均值为 $E[X_n]$ 、方差为 O 的常量分布，即 $\bar{X}_n$ 依概率收敛于 $E[X_n]$. □

Theorem 1.3 (切比雪夫大数定理 (Chebyshev's Law of Large Numbers)). 如果一个随机变量序列 $\{X_n\}$ 独立同分布，且存在常数 $E[X_n] = \mu$ 和 $\mathrm{Var}[X_n] = \sigma^2 < \infty$ ，那么对于任意 $\epsilon > 0$ ，有

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) = 0$$

.

Proof. 这个定理的证明可以直接使用切比雪夫不等式. 对于任意 $\epsilon > 0$ ，我们有

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) \leq \frac{\operatorname{Var}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2},$$

由于 σ^2 是有限的，当 n 趋于无穷时，上式右侧趋于 0. 因此，

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) = 0.$$

□

Theorem 1.4 (马尔科夫大数定理 (Markov's Law of Large Numbers)). 如果一个随机变量序列 $\{X_n\}$ 独立同分布，且存在常数 $E[X_n] = \mu$ ，那么对于任意 $\epsilon > 0$，有

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) = 0.$$

Proof. 我们定义 $S_n = X_1 + X_2 + \ldots + X_n$ 和 $\bar{X}_n = \frac{S_n}{n}$ 。由于 $X_1, X_2, \ldots, X_n$ 是独立同分布的，我们有 $E[X_i] = \mu$ 和 $\operatorname{Var}[X_i] = \sigma^2$ (我们假设方差存在且有限). 我们想要证明的是

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - \mu\right| \geq \epsilon\right) = 0,$$

根据切比雪夫不等式，我们有

$$P\left(\left|\bar{X}_n - \mu\right| \geq \epsilon\right) \leq \frac{\operatorname{Var}\left[\bar{X}_n\right]}{\epsilon^2},$$

由于 $\operatorname{Var}\left[\bar{X}_n\right] = \frac{\sigma^2}{n}$ (这个结果来自方差的性质和随机变量的独立性)，我们可以进一步得到

$$P\left(\left|\bar{X}_n - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2},$$

当 n 趋于无穷大时，上式的右侧趋于 0 ，因此我们得到

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - \mu\right| \geq \epsilon\right) = 0.$$

□