大数定律与中心极限定律

Dezeming Family

2021 年 7 月 7 日

DezemingFamily 系列书和小册子因为是电子书，所以可以很方便地进行修改和重新发布。如果您获得了 DezemingFamily 的系列书，可以从我们的网站 [https://dezeming.top/] 找到最新版。对书的内容建议和出现的错误欢迎在网站留言。

20210708：完成第一版。

20210718：修改了一些符号错误。

目录

一 基本

我们在很多地方推导期望和收敛率时都会直接或者间接使用大数定律或者中心极限定律里的内容，具有很重要的意义。而且可以说，大数定律与中心极限定律是概率论由简单的概率分布扩展到概率理论的过渡。

基本介绍不宜过多，接下来就开始介绍具体的理论。

二 切比雪夫不等式

设随机变量的期望和方差分别为 $E(X)$ 和 $D(X)$，对于任意的 $\epsilon > 0$，都有：

$$P(|X - E(X)| \geq \epsilon) \leq \frac{D(X)}{\epsilon^2} \tag{二.1}$$

这个式子怎么直观理解呢？我们把 $|X - E(X)|$ 作为 X 与期望的距离，该距离大于某值的概率是要小于一定的值的。换言之，该不等式限制了某个随机变量 X 偏离期望的概率。

证明如下，设 X 的概率密度为 $f(x)$：

$$P(|X-E(X)|\geq\epsilon)=\int_{\{|X-E(X)|\geq\epsilon\}}f(x)dx \tag{二.2}$$

$$\leq\int_{\{|X-E(X)|\geq\epsilon\}}\frac{[x-E(X)]^2}{\epsilon^2}f(x)dx \tag{二.3}$$

$$\leq\frac{1}{\epsilon^2}\int_{-\infty}^{+\infty}[x-E(X)]^2f(x)dx \tag{二.4}$$

$$=\frac{D(X)}{\epsilon^2} \tag{二.5}$$

因为 $(|X-E(X)|\geq\epsilon)$ 与 $(|X-E(X)|<\epsilon)$ 互为对立事件，所以：

$$P(|X-E(X)|<\epsilon)\geq 1-\frac{D(X)}{\epsilon^2} \tag{二.6}$$

我们固定 ϵ，可以看到，当方差 $D(X)$ 越小，则 X 取值会越容易贴近于 $X(X)$，因此该不等式可以说明方差可以描述 X 对其均值 $E(X)$ 的偏离程度。

三　大数定律

大数定律用来说明，当采样的样本越多，则平均值就会越接近真的平均值，而且越来越稳定。

设一共有 n 次独立重复试验，其中事件 A 发生了 n_A 次，p_A 是每次试验中事件 A 发生的概率，因此，对于某个大于 0 的 ϵ：

$$\lim_{n\to\infty}P\Big(\Big|\frac{n_A}{n}-p_A\Big|<\epsilon\Big)=1 \tag{三.1}$$

说明随机事件出现的频率具有一定的稳定性，因此可以使用足够多的样本去估计平均值。

而大数定律又分为强大数定律和弱大数定律。

弱大数定律认为，采样次数越多，平均值接近真实的期望值的可能性越大，即我们上面描述的内容：

$$\lim_{n\to\infty}P\Big(\Big|\frac{n_A}{n}-p_A\Big|<\epsilon\Big)=1 \tag{三.2}$$

强大数定律认为，采样次数越多，平均值几乎一定会越来越接近真实的期望值：

$$P\Big(\lim_{n\to\infty}\Big|\frac{n_A}{n}-p_A\Big|<\epsilon\Big)=1 \tag{三.3}$$

看着似乎是一样的，但其实很不相同。我们设求得的样本均值为 $\mu_A=\frac{n_A}{n}$。

弱大数定律中，当 n 趋近于 ∞ 时，μ_A 接近真实的 p_A 的可能性会越来越大，朝着 p_A 去的概率越来越大，但是也有概率偏离 p_A。

强大数定律中，当 n 趋近于 ∞，μ_A 几乎一定能不断的接近 p_A，换言之，μ_A 不断朝着 p_A 的方向过渡。

四　中心极限定律

中心极限定理稍微有点绕，但基本含义也不是很难。中心极限定理可以理解为，不论变量在总体中的分布如何，当样本量足够大时，变量均值的采样分布会近似于正态分布。

公式描述为：设样本期望 $E(X_k)=\mu$，方差为 $D(X_k)=\sigma^2$，$X_1,X_2,...,X_n$ 是相互独立且相同分布的随机变量：

$$\lim_{n\to\infty}P\Big(\frac{\sum_{k=1}^{n}X_k-n\mu}{\sigma\sqrt{n}}\leq x\Big)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x) \tag{四.1}$$

我们把中间的式子稍微转换一下：

$$\frac{\sum_{k=1}^{n}X_k-n\mu}{\sigma\sqrt{n}}=\frac{\frac{1}{n}\sum_{k=1}^{n}X_k-\mu}{\sigma/\sqrt{n}} \tag{四.2}$$

设 $\frac{1}{n}\sum_{k=1}^{n} X_k = \overline{X}$，因此我们可以简化中心极限定理为：

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N\big(0,1\big) \tag{四.3}$$

$$\bar{X} \sim N\big(\mu,\frac{\sigma^2}{n}\big) \tag{四.4}$$

其中，$\sim$ 表示为“近似服从”。

我们来用具体例子解释一下，我们有一个变量，符合任意分布，然后我们独立随机采样 n 个，计算出它的均值 $\bar{X}_1$，然后我们再独立采样 n 个，计算出均值 $\bar{X}_2$，以此类推，我们一共采样了 100 次，然后我们发现这些均值 $\bar{X}_k$ 很像正态分布！而且当 n 的值越大，就越像正态分布，n 无穷大时，样本均值分布即符合正态分布。

参考文献

[1] 吴臻, 刘建亚. 概率论与数理统计 [M]. 山东大学出版社, 2004.

[2] https://www.zhihu.com/question/21110761