Lecture 05: Random Variable and Distribution Function

Definitions

$\boxed{\text{Definition 5.1}}$ 设 $X:\Omega\to\mathbb{R}$，且对于任意 $\mathbb{R}$ 中的 Borel 集 B，$\{e|X(e)\in B\}\in\mathscr{F}$，则称 X 是 $(\Omega,\mathscr{F},P)$ 上的随机变量。

注：1. Borel 集是由所有的 $\{(a,b]|-\infty\le a,b\le+\infty\}$ 经过可数次交或并得到的集合。

2. 在大部分场合，只需关注 X 是从 Ω 到 $\mathbb{R}$ 上的映射即可。

【例】示性随机变量 (indicator)，对于 $A\in\mathscr{F}$，定义 $X_A(e)=\begin{cases}1,e\in A\\0,e\notin A\end{cases}$。

$\boxed{\text{Definition 5.2}}$ 设 X 是随机变量，则称 $F_X(x)\overset{def}{=}P(X\le x)$ 为 X 的分布函数。

分布函数满足以下性质：

- 单调不降。

  > 对于 $x_1<x_2$，$F(x_2)-F(x_1)=P(x_1<x\le x_2)\ge 0$。

- $\lim_{x\to+\infty}F(x)=1,\lim_{x\to-\infty}F(x)=0$。

  > Lemma (单调收敛定理) 当被积函数单调递增时，积分和极限可以换序。

  > $\{X\le x\}\xrightarrow{x\to+\infty}\Omega$，$lim_{x\to\infty}P(X\le x)\overset{lemma}{=}P(\lim_{x\to\infty}\{X\le x\})=P(\Omega)=1$。

- F 右连续且存在左极限 (càdlàg, RCLL)，i.e. $\lim_{x\to x_0^+}F(x)=F(x_0)$，$F(x_0-0)$ 存在。

  > $F(x)-F(x_0)=P(X\le x)-P(X\le x_0)=P(x_0<X\le x)\xrightarrow{x\to x_0^+}P(\emptyset)=0$。
  > $x<x_0$ 时，$F(x)\le F(x_0)$ 且 $F(x)$ 单增，所以左极限存在。

注：左不一定连续的原因是：$F(x_0)-F(x)=P(x<X\le x_0)\xrightarrow{x\to x_0^-}P(X=x_0)$ 不一定为 0。

$\boxed{\text{Theorem 5.3}}$ 如果一个函数 F 满足上述三条性质，那么它必然是某个随机变量 X 的分布函数。

Discrete Random Variable

$\boxed{\text{Definition 5.4}}$ 若随机变量 X 的取值为有限多个或无限可数个，则 X 为离散随机变量。设 X 的取值为 $x_1,x_2,\cdots$，令 $P(x=x_k)=P_k$，称 $\{P_k\}$ 为 X 的分布列/分布律。

注：$F(x)=P(X\le x)=P(\bigcup_{x_k\le x}P_k)=\sum_{x_k\le x}P_k$。

Common Distributions

【例】(0-1分布) 设 $A\in\mathscr{F}$，$P(A)=p\in(0,1)$。令 X_A 为 A 的 indicator，则 $P(X_A=1)=P(A)=p,P(X_A=0)=1-p$。

【例】(二项分布) 设 X 的分布律为 $p_k=P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},p\in(0,1)$，$p_k$ 即为 n 重 Bernoulli 试验成功 k 次的概率，X 服从二项分布，记为 $X\sim B(n,p)$。

【例】(泊松分布) 设 X 的分布律为 $p_k=P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$，则称 X 服从泊松分布，记为 $X\sim P(\lambda)$。

> $$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\overset{Taylor\ Series}{=}e^{-\lambda}e^{\lambda}=1$$

注：在 $\lambda=np_n$ 固定的情况下，当 n 很大时，p_n 则很小。那么在 $k<<n$ 的情况下，二项分布可以近似为泊松分布。

【例】(几何分布) 独立重复试验，成功概率为 p，记第 k 次试验首次成功的概率为 $p_k=P(X=k)=(1-p)^{k-1}p$，则 X 服从几何分布，记为 $X\sim g(p)$。

> $$\sum_{k=1}^{\infty}(1-p)^{k-1}p=p\sum_{k=0}^{\infty}(1-p)^k=p\cdot\frac{1}{1-(1-p)}=p\cdot\frac{1}{p}=1$$

几何分布没有记忆性。假设前 t 次试验均失败，则再试 s 次成功的概率 $P(X=t+s|X>t)=P(X=s)$。或者：$P(X\geq t+s|x>t)=\sum_{k=s}^{\infty}P(X=t+k|X>t)=\sum_{k=s}^{\infty}P(X=k)=P(X\geq s)$。

【例】(巴斯卡分布) 独立重复试验，每次成功概率为 p，记做完第 k 次试验后恰好取得了 r 次成功 $(k\geq r)$ 的概率为 p_k，有

$$p_k=P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r}$$

我们称 X 服从巴斯卡分布 ($r=1$ 时的巴斯卡分布就是几何分布)。

> $$\sum_{k=r}^{\infty}p_k=\sum_{k=r}^{\infty}\binom{k-1}{r-1}p^r(1-p)^{k-r}=p^r\sum_{k=r}^{\infty}\binom{k-1}{r-1}(1-p)^{k-r}$$
>
> 令 $q=1-p$，只要证 $\sum_{k=r}^{\infty}\binom{k-1}{r-1}q^{k-r}=(1-q)^{-r}$。对 $f(x)=(1-x)^{-r}$ 泰勒展开，
>
> $$\begin{aligned}f(x)=(1-x)^{-r}&=\sum_{k=0}^{\infty}\frac{f^{(k)}(0)}{k!}x^k\\&=\sum_{k=0}^{\infty}\frac{1}{k!}(-1)^k(-r)(-r-1)\cdots(-r-k+1)x^k\\&=\sum_{k=0}^{\infty}\frac{1}{k!}r(r+1)\cdots(r+k-1)x^k\\&=\sum_{k=0}^{\infty}\frac{(r+k-1)!}{k!(r-1)!}x^k\\&=\sum_{k=0}^{\infty}\binom{r+k-1}{k}x^k\\&=\sum_{k=r}^{\infty}\binom{k-1}{r-1}x^{k-r}\end{aligned}$$

【例】(超几何分布) 对 N 个产品无放回抽样 n 次，其中有 M 个次品，令抽到 k 个次品的概率为 p_k，则

$$p_k=P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

我们称 X 服从超几何分布。

> $$\sum_{k=0}^{\infty} p_k = \frac{1}{\binom{N}{n}} \sum_{k=max(0,n-(N-M))}^{min(n,M)} \binom{M}{k}\binom{N-M}{n-k}$$
>
> 右边的部分用“讲故事法”容易证明等于 $\binom{N}{n}$。如果需要比较数学的方法，可以从 $(1+x)^N$ 和 $(1+x)^M(1+x)^{N-M}$ 两个角度去考察 x^n 前的系数。

注：当 $N, N-M >> n \geq k$ 时，超几何分布可以近似地当作二项分布计算：

$$\begin{aligned}\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} &= \frac{n!}{k!(n-k)!} \frac{M(M-1)\cdots(M-k+1)}{N(N-1)\cdots(N-k+1)} \frac{(N-M)\cdots(N-M-(n-k)+1)}{(N-k)\cdots(N-k-(n-k)+1)} \\ &\approx \binom{n}{k}\left(\frac{M}{N}\right)^k\left(\frac{N-M}{N}\right)^{n-k}\end{aligned}$$