2.2 常见分布

2.2.1 二项分布

定义 2.6 设每次试验中事件 A 发生的概率为 p。n 重伯努利试验中，事件 A 发生的次数设为 X，X 可能取值为 $0,1,2,3\ldots n$, $P(X=k)=C_k^n\times p^k\times q^{n-k}$，其中 $q=1-p, 0<p<1$, $k=0,1,2,3\ldots n$ 则称随机变量 X 服从参数为 n, p 的二项分布, 记为 $X\sim B(n,p)$。

推导： n 重伯努利试验，$\Omega=\{\omega=(\omega_1,\omega_2\cdots\omega_n)|\omega_i=0/1, 1\leq i\leq n\}$。事件 $\{X=k\}=\{\omega|\omega$ 的取值有 k 个 1，$n-k$ 个 0$\}$，有 $P(\omega)=p^k(1-p)^{n-k}, \forall\omega\in\{X=k\}$。事件 {X=k} 中的 ω 有 C_k^n 个，且互不相交，所以 $P(X=k)=C_k^n\times p^k\times q^{n-k}$，其中 $q=1-p, 0<p<1$, $k=0,1,2,3\ldots n$

当 n=1 时, 这就是（0-1）分布，所以（0-1）分布是二项分布的特例。

$E(X)=np$

$Var(X)=np(1-p)$

2.2.2 泊松分布

定义 2.7 若随机变量 X 的分布律为:$P(X=K)=\frac{\lambda^k}{k!}\times e^{-\lambda}$，$\lambda>0$, $k=0,1,2,3\ldots n$, 则称随机变量 X 服从参数为 λ 的泊松分布, 记为 $X\sim Poisson(\lambda)$。

泊松分布为二项分布的极限分布（$n\times p=\lambda$，$n\to\infty$）。

解释： 设有一系列服从二项分布的随机变量 X_n, $X_n\sim B(n,p)$, 若 $\lim_{n\to+\infty} n\times p_n=\lambda$，$n$ 很大，p 很小，np 大小适中时，$\lim_{n\to+\infty} P(X_n=k)=C_k^n\times p^k\times q^{n-k}=\frac{\lambda^k}{k!}\times e^{-\lambda}$, 即泊松分布为二项分布的极限分布。

例 2.5 （*1*）某一时间内，电话总机接到用户呼唤的次数
（*2*）某一天进入紫荆超市的顾客数目

$E(X)=\lambda$；$Var(X)=\lambda$

2.2.3 几何分布

定义 2.8 无限重伯努利试验中，每次成功的概率为 p，记 X 为首次出现成功时的试验次数，*X=1*，*2*，*3*...*n*，若随机变量 X 的分布律为 $P(X=k)=q^{k-1}\times p$, 其中 $p\geq 0$，$q=1-p$, 则称随机变量 X 服从参数为 p 的几何分布，记为 $X\sim G(p)$。

推导： $\{X=k\}$=“首次成功在第 k 次”= 第 *1* 次失败 $\bigcap$ 第 *2* 次失败 $\bigcap\cdots$ 第 $(k-1)$ 次失败 $\bigcap$ 第 k 次成功，所以 $P(X=k)=q^{k-1}\times p$，$k=1,2\cdots n$

验证： $\sum_{k=1}^{+\infty}(1-p)^{k-1}\times p=1$

定义 2.9 无记忆性：$P(X>m+n|X>m)=P(X>n)$，$\forall m,n\in N$,

注： 前 m 次试验失败的条件下，在接下来的 n 次试验中，仍然失败的概率只与 n 有关，而与之前失败的 m 次试验无关，似乎忘记了均失败的前 m 次试验结果，这就是无记忆性。

推导： $P(X>n)=\sum_{k=n+1}^{+\infty}(1-p)^{k-1}p=(1-p)^n$

$P(X>m+n|X>m)=\frac{P(X>m+n)}{P(X>m)}=\frac{(1-p)^{m+n}}{(1-p)^m}=(1-p)^n$

2.2.4 指数分布

定义 2.10 设随机变量 X 的密度函数为 *f(x)*, 若 *f(x)* 满足：

$$f(x)=\begin{cases}\lambda\times e^{-\lambda\times x} & x\geq 0\\ 0 & x<0\end{cases}$$

$\lambda>0$, 则称随机变量 X 服从参数为 λ 的指数分布。

$E(X)=\frac{1}{\lambda};Var(X)=\frac{1}{\lambda^2}$

在实践中，指数分布经常作为寿命，等待时长，或者间隔时长的分布而出现。指数分布是唯一具有无记忆性的连续分布，无记忆性是指数分布特有的一个性质。

例 2.6 从时间零点开始，你在二校门等你的朋友，假设你的朋友的到达时间 X 服从参数为 λ 的指数分布，过了 s 小时，你的朋友没有来，那么你再等 t 小时，你的朋友仍没有来的概率只与时间 t 有关，而与你已经等待的 s 小时无关，似乎已经忘记了让你已经等待过的 s 小时，这就是指数分布的无记忆性。

证明： $\forall s\ t\geq 0$，$P(X>t)=e^{-\lambda\times t}$，（$t\geq 0$）

左边 = $\frac{P(X>s+t)}{P(X>s)}=\frac{e^{-\lambda\times(s+t)}}{e^{-\lambda\times s}}=e^{-\lambda\times t}=P(X>t)$,

$P(X>s+t|X>s)=P(X>t)$

例 2.7 某邮局有两个职员，假设当 c 走进邮局时，发现两个职员正在接待 a 和 b，于是被告知一旦处理完 a 或者 b 的事情，就立即接待他，如果每位顾客的服务时长服从参数为 λ 的指数分布，那么 c 是最后一个办完事情的概率有多大？

解： 从 c 开始接受服务时开始考虑（零点选择）。此时，a 和 b 有一个已经离开，另一个仍在继续接受服务，然而因为指数分布具有无记忆性，仍在接受服务的顾客（a 或者 b）的办事时长仍然服从参数为 λ 的指数分布，这与此时刚开始服务是一样的。因此由对称性得，剩下的一个人在 c 前完成服务的概率为 *0.5.*

2.2.5 正态分布

定义 2.11 如果随机变量 X 的概率密度为 $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < +\infty$，其中，$\mu, \sigma(\sigma > 0)$ 为常数，则称 X 服从参数为 μ, σ 的正态分布（或高斯分布），记为 $X \sim N(\mu, \sigma^2)$.

$E(X) = \mu; Var(X) = \sigma^2$

当 $\mu = 0, \sigma = 1$ 时，即 $X \sim N(0\,,1)$ 时，称 X 服从标准正态分布，其密度函数为：$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, -\infty < x < +\infty$