定理 2.1.3 设超几何分布中, D 是 N 的函数, 且满足

$$\lim_{N\to\infty}\frac{D}{N}=p.$$

其中 $0<p<1$. 固定 n. 那么对任意 $k\geqslant 0$,

$$\lim_{N\to\infty}\frac{C_D^k C_{N-D}^{n-k}}{C_N^n}=C_n^k p^k(1-p)^{n-k}.$$

- 由假设 $N\to\infty$ 时, $D\to\infty$. 不妨认为 $n<D<N$.

$$\frac{C_D^k C_{N-D}^{n-k}}{C_N^n}$$

$$
\begin{aligned}
&= \frac{\dfrac{D!}{k!\,(D-k)!}\,\dfrac{(N-D)!}{(n-k)!\,(N-D-n+k)!}}{\dfrac{N!}{n!\,(N-n)!}} \\
&= \frac{n!}{k!\,(n-k)!}\,\frac{D!}{(D-k)!}\,\frac{(N-D)!}{(N-D-n+k)!}\,\frac{(N-n)!}{N!} \\
&= C_n^k \frac{D\cdots(D-k+1)}{N^k}\,\frac{(N-D)\cdot\cdots\cdot(N-D-n+k+1)}{N^{n-k}} \\
&\frac{N^n}{N\cdots(N-n+1)} \\
&= C_n^k \prod_{i=1}^{k}\left(\frac{D}{N}-\frac{i-1}{N}\right)\prod_{i=1}^{n-k}\left(1-\frac{D+i-1}{N}\right)\prod_{i=1}^{n}\frac{N}{N-i+1} \\
&\to C_n^k p^k (1-p)^{n-k}\,.
\end{aligned}
$$

例题 2.1.4 某地区有很多猴子, 但具体数量 N 无法得知, 请设计实验方法对猴子的数量 N 进行估计.

- 任取 m 只猴子, 做上标记后放回. 之后再任取 n 只猴子, 其中带有标记的猴子数量为随机变量 X, 设检查发现带有标记的有 k 只, $k \in \{0, 1, ..., \min(n, m)\}$. 利用超几何分布

$$P_N(X = k) = \frac{C_m^k C_{N-m}^{n-k}}{C_N^n}.$$

注意, $i < j$ 时,

$$C_j^i = \frac{j!}{i!(j-i)!} = \frac{j}{j-i} \frac{(j-1)!}{i!(j-i-1)!} = \frac{j}{j-i} C_{j-1}^i$$

因此

$$\begin{aligned}\frac{P_N(X=k)}{P_{N-1}(X=k)} &= \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} \frac{C_{N-1}^n}{C_m^k C_{N-1-m}^{n-k}} \\ &= \frac{C_{N-1}^n}{C_N^n} \frac{C_{N-m}^{n-k}}{C_{N-1-m}^{n-k}} \\ &= \frac{N-n}{N} \frac{N-m}{N-m-n+k} > 1,\end{aligned}$$

化简得

$$N^2 - mN - nN + mn > N^2 - mN - nN + Nk,$$

即

$$N < \frac{mn}{k}.$$

故

$$N = \left[\frac{mn}{k}\right].$$

几何分布

重复某一实验, 直到实验成功为止. 每次实验成功率为 p, $0 < p < 1$. 至首次成功时的试验次数 X 是一个随机变量, 其分布称为几何分布. 显然 ,

$$P\left(X = k\right) = \left(1 - p\right)^{k-1} p,\ k \in \{1, 2, ...\}.$$

连续分布

定义 2.1.6 随机变量 X 称为**连续分布**, 如果存在定义在 $\mathbb{R}$ 上的非负函数 $f(x)$ 使得下面的等式成立,

$$P(X \in E) = \int_E f(x)\,dx$$

其中 E 是任意有界, 无界, 开或者闭的实数区间. **非负**函数 $f(x)$ 则称为该连续分布的**密度函数**.

注记 2.1.1 定义并不假定密度函数 $f(x)$ 的连续性, 或有界性.

定理 2.1.4 由概率公理可知, 密度函数应满足

$$\int_{\mathbb{R}} f(x)\,dx = 1.$$

定理 2.1.5 假设 X 是具有密度函数 $f(x)$. 那么对任意实数 $a \in \mathbb{R}$,

$$P(X = a) = 0$$

因此, $\forall b > a$ 总有 $P(a < X < b) = P(a < X \leqslant b)$.

从该定理可知, 对一个连续型随机变量, 其分布的密度函数并不是唯一的, 例如我们可以修改密度函数在有限个点上的值而不改变原来的分布函数. 对应于同一个分布函数的密度函数都是几乎处处相等的.

定义 2.1.7 给定两个函数 f, g 以及事件 $D = \{f \neq g\}$. 如果 $P(D) = 0$, 那么我们称 f 和 g 是**几乎处处**相等的, 写作 $f = g$, $a.e.$

例题 2.1.5 如下函数满足密度函数的条件

$$f(x)=\begin{cases} 0, & x \leqslant 0 \\ \dfrac{1}{(1+x)^2}, & x>0 \end{cases}$$

例题 2.1.6 如下无界函数满足密度函数的条件

$$f(x)=\begin{cases} \dfrac{2}{3}x^{-1/3}, & 0<x<1 \\ 0, & \text{其他} \end{cases}$$

一致分布（均匀分布）

定义 2.1.8 如果随机变量 X 的取值在区间 $[a,b]$, 并且 X 取值于 $[a,b]$ 的任何子区间的概率与子区间长度成比例. 那么随机变量 X 称为在 $[a,b]$ 上具有一致分布.

定理 2.1.6 如果随机变量 X 称为在上具有一致分布, 那么 X 的密度函数是

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leqslant x \leqslant b \\ 0, & \text{其他} \end{cases}$$

■ 由于 X 取值在区间 $[a, b]$ 上, 其密度函数在此区间之外都应为 0. 任取 $x \in (a, b)$, $\epsilon > 0$ 使得 $[x-\epsilon, x+\epsilon] \subset [a, b]$. 根据一致分布的定义, 等式

$$P\left(x-\epsilon \leqslant X \leqslant x+\epsilon\right) = \int_{x-\epsilon}^{x+\epsilon} f(t)\, dt$$

不依赖于 x 的位置, 对 x 求导数得到

$$0 = \frac{d}{dx}\int_{x-\epsilon}^{x+\epsilon} f(t)\,dt = f(x+\epsilon) - f(x-\epsilon)\,,\ a.e.$$

由 x 和 ϵ 的任意性, 可知 f 在区间 $[a,b]$ 上为常数（$a.e.$）. 运用

$$\int_a^b f(x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx = 1$$

即可得出定理结论.

2.2　随机变量的分布函数

定义 2.2.1 给定概率空间以及概率 P. 该空间上的随机变量 X 的 **(累积) 分布函数**定义为

$$F(x) = P(X \leqslant x), \forall x.$$

由概率的基本性质可以得出分布函数的以下性质

定理 2.2.1 对于分布函数成立

(i) $P(x_1 < X \leqslant x_2) = F(x_2) - F(x_1)$;

(ii) $x_1 \leqslant x_2 \Longrightarrow F(x_1) \leqslant F(x_2)$, 即 F 是单调不减函数.;

(iii) $F(-\infty) = 0$, $F(\infty) = 1$;

(iv) $F(x) = F(x+)$, 即 F 是右连续函数.

(v) $P(X < x) = F(x-)$;

(vi) $P(X=x)=F(x)-F(x-)$;

■ (i) 只需注意 $\{X \leqslant x_2\}$ 可写成不相交集合的并集:

$$\{x_1 < X \leqslant x_2\} \cup \{X \leqslant x_1\}.$$

(ii) $\{X \leqslant x_1\} \subset \{X \leqslant x_2\}$

(iii) $F(-\infty)$ 的含义是 $F(-\infty)=\lim_{x\to-\infty} P(X \leqslant x)$, 然而 $\{X \leqslant x\}$ 是关于 x 的不减少事件, 当 $x \to -\infty$ 时 $\{X \leqslant x\}$ 收敛到空集, 因此由概率的上连续性 (参见上一课笔记) 可知 $F(-\infty)=0$. 同理可证 $F(\infty)=0$.

(iv) 根据上一性质, $F(x)$ 存在左右极限. 由于 $\{X \leqslant y\}$ 是关于 y 的不减少事件, 当 $y \downarrow x$ 时, 该集合下降收敛到 $\{X \leqslant x\}$, 由概率的上连续性得到 $\lim_{y\downarrow x} F(y)=F(x)$, 正是右连续性.

(v) 当 $y \uparrow x$ 时, $\{X \leqslant y\}$ 递增收敛到 $\{X < x\}$, 由概率的下连续性,

$$F(x-) = \lim_{y \uparrow x} F(y) = P(X < x).$$

(vi) 注意到 $\{X \leqslant x\} = \{X = x\} \cup \{X < x\}$.

如果一个函数 $H(x)$ 满足以下条件:(1) $H(x) \geqslant 0$, $H(x)$ 单调不减; (2) $H(x)$ 右连续; (3) $H(-\infty) = 0$, $H(\infty) = 1$; 那么存在随机变量 X, 使得 X 的分布函数正好是 $H(x)$, 因此我们通常把满足上面三个条件的函数直接地称为**分布函数**, 而不指明与其对应的随机变量.

定理 2.2.2 假设连续型随机变量 X 的密度函数是 $f(x)$, 那么它

的分布函数

$$F(x) = \int_{-\infty}^{x} f(z)\,dz$$

是连续的, 并且 $F(x)$ 在 $f(x)$ 的连续点上导数存在, 即, 如果 $f(x)$ 在 x 连续那么

$$F'(x) = f(x)$$

- 连续随机变量在任何单点处概率为 0, 因此 $F(x) - F(x-) = 0$, 因此是左连续的, 从而连续. $\forall x, h \neq 0$,

$$\begin{aligned} \frac{F(x+h) - F(x)}{h} - f(x) &= \frac{1}{h}\int_{x}^{x+h} f(z)\,dz - f(x) \\ &= \frac{1}{h}\int_{x}^{x+h} \left(f(z) - f(x)\right) dz. \end{aligned}$$

由于 $f(z)$ 在 x 处连续, $\forall\varepsilon>0$, 存在 $\delta>0$, 当 $|z-x|<\delta$ 时有

$$|f(z)-f(x)|<\varepsilon.$$

从而当 $|h|<\delta$ 时

$$\left|\frac{F(x+h)-F(x)}{h}-f(x)\right|\leqslant\frac{1}{h}\int_x^{x+h}|f(z)-f(x)|\,dz\leqslant\varepsilon,$$

这表明

$$\lim_{h\to0}\frac{F(x+h)-F(x)}{h}=f(x),$$

即

$$F'(x)=f(x).$$

例题 2.2.1 假设 X 是一个参数 p 的 Bernoulli 随机变量, 确定其分布函数.

■ 根据定义 $P(X=1)=p, P(X=0)=1-p$. 因此其分布函数为

$$F(x)=\begin{cases} 0, & x<0; \\ 1-p, & 0\leqslant x<1; \\ 1 & x\geqslant 1. \end{cases}$$

例题 2.2.2 假设随机变量 X 在 $[a,b]$ 上具有一致分布, 确定其分布函数.

- 其分布函数为

$$F(x)=\begin{cases}0, & x\leqslant a;\\ \displaystyle\int_a^x \frac{1}{b-a}du=\frac{x-a}{b-a}, & a<x\leqslant b;\\ 1 & x>b.\end{cases}$$

例题 2.2.3 验证下面的函数是一个分布函数, 并确定其密度函数,

$$F(x)=\begin{cases}0, & x\leqslant 0\\ 1-\dfrac{1}{1+x}, & x>0\end{cases}.$$

- 容易验证 $F(x)$ 是一个分布函数. 除了 $x=0$ 以外, $F(x)$ 的

导数处处存在, 因此其密度函数可以写作

$$f(x)=\begin{cases}0, & x\leqslant 0\\ \dfrac{1}{(1+x)^2}, & x>0\end{cases}.$$

指数分布

回到 Poisson 分布一节银行客户到来数量问题. Poisson 分布已经给出了客户到来的数量的分布, 现在银行还想知道客户到来的时间间隔分布情况. 由于客户到来是相互独立的, 因此各个客户到来之间的时间间隔同分布, 我们只需知道首个客户的到来时间 T 的分布情况即可. 为此将时间数量化, 假设客户只能在时间点 $0, \delta, 2\delta,\dots$ 上以概率 p_δ 到来. 于是任意给

定时刻 t 都包括在某小时间区间内: $n\delta \leqslant t < (n+1)\,\delta$, 近似地就有

$$P\,(T > t) \approx P\,(T > n\delta) = (1 - p_\delta)^n$$

由期望计算公式

$$ET = \sum_{n=0} P\,(T > n\delta)\,\delta = \frac{\delta}{p_\delta}$$

这提示在极限情况 $(\delta \to 0)$ 下, 我们应合理地假设, 存在常数 $\alpha > 0$ 使得 $\delta/p_\delta = \alpha$. 这样 (注意 $n\delta \to t$)

$$P\,(T > t) \approx (1 - p_\delta)^n \approx \exp\,(-np_\delta) = \exp\left(-n\delta \cdot \frac{p_\delta}{\delta}\right) \approx e^{-t/\alpha}$$

容易看到

$$1 - P\,(T > t) \approx 1 - e^{-t/\alpha}$$

是一个分布函数, 这一极限情形下的分布就称为指数分布.

定义 2.2.2 $\beta > 0$. 随机变量 T 称为具有参数 β 的**指数分布**, 如果 T 的密度函数为

$$f(t|\beta) = \begin{cases} \beta e^{-\beta t}, & t > 0 \\ 0, & t \leqslant 0 \end{cases}$$

指数分布可视为 Gamma 分布的一个特殊情形.

灯泡的使用寿命 T 服从参数 β 的指数分布 $Exp(\beta)$. 一人走进房间发现灯泡亮着, 用 X 表示从此刻起灯泡还能继续使用的时间. 试确定 X 的分布.

■ 假设从灯泡点亮到当此人走进房间相差的时间为 $t > 0$, 那么要求的是在 $T > t$ 的条件下 X 的分布, 即

$$P(X > x + t | T > t), \ \forall x > 0.$$

根据指数分布无记忆性, 这仍然是参数 β 的指数分布. ///

定理 2.2.3 (**指数分布无记忆性**) 随机变量 T 服从参数为 β 的指数分布. 那么对任意 $t > 0$ 都有

$$P(T > x + t | T > t) = P(T > x)$$

- $\forall x > 0$,

$$P(T > x) = \int_x^{\infty} \beta e^{-\beta t} dt = e^{-\beta x}$$

因此

$$P(T > x + t | T > t) = \frac{P(T > x + t)}{P(T > t)} = \frac{e^{-\beta(x+t)}}{e^{-\beta t}} = e^{-\beta x}.$$

正态分布

随机变量 X 具有参数为 $\mu \in \mathbb{R}, \sigma^2$ $(\sigma > 0)$ 的正态分布是指它的密度函数为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \ \forall x \in \mathbb{R}.$$

记号 $X \sim N\left(\mu, \sigma^2\right)$.

容易计算 $\mu = 0, \sigma^2 = 1$ 时,

$$\begin{aligned}
&\left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{x^2}{2}\right) dx\right)^2 \\
= &\frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(x^2 + y^2\right)\right) dxdy
\end{aligned}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \left(\int_0^{\infty} \exp\left(-\frac{1}{2} r^2 \right) r dr \right) d\theta = 1.$$

因此

$$\int_{\mathbb{R}} p(x)\, dx = 1.$$

由此

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{y^2}{2} \right) dy \ \left(y = \frac{x-\mu}{\sigma} \right) \\
&= 1
\end{aligned}$$