例1.2. 盒中有5 个球, 其中有2 个白球, 3 个黑球.
从中任取3 个球, 将其中所含的白球的数目记为X.

- 建模: 将球编号, $1 \sim 3$ 表示黑球, $4, 5$ 表示白球.
- $\omega = (i, j, k)$, 其中$1 \leqslant i < j < k \leqslant 5$. $\Omega = C_5^3 = 10$.
- 满足$X = 0$ 的ω 有$C_2^0 C_3^3 = 1$ 个;
  满足$X = 1$ 的ω 有$C_2^1 C_3^2 = 6$ 个;
  满足$X = 2$ 的ω 有$C_2^2 C_3^1 = 3$ 个.
- 事件: $\{X = 1\} = \{\omega : X(\omega) = 1\}$,
  $$\{X \leqslant 1\} = \{\omega : X(\omega) \leqslant 1\}.$$
- 将$P(\{X = 1\})$ 简记为$P(X = 1)$. 例如,
  $$P(X = 1) = \frac{6}{10},\ P(X \leqslant 1) = \frac{7}{10}.$$

例1.6. 某公共汽车站每隔10 min 会有一辆某路公交车到达. 某乘客随机在任意时刻到达车站.

- 候车时间X (单位: min) 为随机变量.
- $0 \leqslant X \leqslant 10$.
- 几何概型(参阅§1.8): 例,

$$P(X \leqslant 3) = \frac{3}{10}, \quad P(2 \leqslant X \leqslant 6) = \frac{4}{10}.$$

§2.2 离散型随机变量

- 定义2.1. X 是离散型随机变量指: X 取有限个值$x_1, \cdots, x_n$, 或可列个值$x_1, x_2, \cdots$. X 的概率分布(列)指:

$$p_k = P(X = x_k), \quad k = 1, \cdots, n \text{ 或 } k = 1, 2, \cdots.$$

- 概率分布表:

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p	p_1	p_2	$\cdots$	p_k	$\cdots$

- 非负: $p_k \geqslant 0, \forall k$; 规范: $\sum_{k=1}^{n} p_k = 1$ 或$\sum_{k=1}^{\infty} p_k = 1$.

1. 两点分布(伯努利分布), $X \sim B(1,p)$ (参数$0 \leqslant p \leqslant 1$):

$$P(X=1)=p, \quad P(X=0)=1-p.$$

- 模型: 投币,

$$\text{投到}H \text{ 则}X=1; \text{投到}T \text{ 则}X=0.$$

- 示性函数1_A:

$$\text{事件}A \text{ 发生则取}1; A \text{ 不发生则取}0.$$

- 例2.1. 100 件产品中有3 件次品. 从中任取一件.

$$A=\text{“取到合格品”}, X=1_A, p=0.97.$$

- $p=0$ 时, $X \equiv 0$; $p=1$ 时$X \equiv 1$. 即, X 退化.

2. 二项分布, $X \sim B(n,p)$ (参数$n \geqslant 1,\ 0 \leqslant p \leqslant 1$):

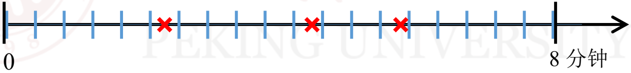
$$P(X=k) = C_n^k p^k (1-p)^{n-k},\ \ k = 0, 1, \cdots, n.$$

- 模型: 独立投币n 次, 正面的总次数.
- 定理2.1. 分布列的最大值点k_0 如下:
  若$(n+1)p \notin \mathbb{Z}$, 则$k_0 = [(n+1)p]$;
  若$(n+1)p \in \mathbb{Z}$, 则$k_0 = (n+1)p$ 或$(n+1)p - 1$.

3. 泊松分布, $X \sim \mathcal{P}(\lambda)$ (参数: $\lambda > 0$):

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \cdots .$$

- 模型: 例2.3. 研究放射性物质在8 分钟内放射出的粒子数X.
- 近似: 以$\frac{1}{n} *$ 8分钟为一个微观时间.



- (i) 在一个微观时间内放射粒子的概率为$p = \frac{\lambda}{n}$,
  (ii) 不同的微观时间内是否放射粒子相互独立.

- X 近似服从$B(n,p)$, 故

$$\begin{aligned} P(X=k) &\approx C_n^k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^n \\ &\approx \frac{(np)^k}{k!}(1-\frac{\lambda}{n})^n = \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

- 上式即为§1.7 第一近似公式.
- 定理2.2. 分布列的最大值点$k_0 \approx \lambda$.

4. 超几何分布, $X \sim H(N, D, n)$ (参数N, D, n):

$$P(X=k)=\frac{C_D^k C_{N-D}^{n-k}}{C_N^n}, \quad k=0,1,\cdots,n.$$

- 模型: N个产品, D 个次品, 任取n 个, 抽到的次品数为X.
- 放回抽样vs 不放回抽样: 二项分布vs 超几何分布.
- 定理2.3. 给定n. 当$N \to \infty$, $\frac{D}{N} \to p$ 时,

$$\frac{C_D^k C_{N-D}^{n-k}}{C_N^n} \to C_n^k p^k (1-p)^{n-k}, \quad k \geqslant 0.$$

- 读解定理2.3: 例$n=5$. H 表示次品, T 表示合格品, 则

$$HHTHT: \frac{D}{N} \cdot \frac{D-1}{N-1} \cdot \frac{N-D}{N-2} \cdot \frac{D-2}{N-3} \cdot \frac{N-D-1}{N-4} \to p^3 q^2.$$

5. 几何分布, $X \sim G(p)$, 参数$0 < p < 1$:

$$P(X = k) = (1 - p)^{k-1}p, \quad k = 1, 2, \cdots .$$

- 模型: 独立重复投币中, 第一次投到H 时的投币次数.
- $P(X > n) = (1 - p)^n, \forall n \geqslant 0.$
- 无记忆性: $P(X - n = k | X > n) = P(X = k).$

- 6. 负二项分布, $X \sim NB(r,p)$, 参数$r \geqslant 1$, $0 < p < 1$:

$$P(X=k) = C_{k-1}^{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \cdots$$

- 模型: 独立重复投币中, 第r 次投到H 时的投币次数.
- 7. 离散均匀分布,

$$P(X=k) = \frac{1}{N}, \quad k = 1, \cdots, N.$$

- 模型: 古典概型.

§2.3 连续型随机变量

- 定义3.1. 连续型随机变量指: 存在$p(x)$ 使得

$$P(a \leqslant X \leqslant b) = \int_a^b p(x)dx, \quad \forall a < b.$$

称$p(\cdot)$ 为X 的概率密度(函数), 也记为$p_X(\cdot)$.

- 非负: $p(x) \geqslant 0$; 规范: $\int_{-\infty}^{\infty} p(x)dx = 1$.
- $P(X = x) = 0$ vs $p(x) \geqslant 0$.
- $p(\cdot)$ 在x 连续, 则$P(X \in [x, x+\Delta x]) = p(x)\Delta x + o(\Delta x)$,
- 单独谈论一个点x 对应的$p(x)$ 没有意义.

1. 均匀分布, $X \sim U(a,b)$ (参数$a<b$):

$$p(x)=\begin{cases}\frac{1}{b-a}; & 若 a \leqslant x \leqslant b;\\ 0, & 否则.\end{cases}$$

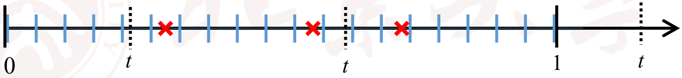
- $a \leqslant x \leqslant b$ 可改为$a<x<b$, $a<x\leqslant b$, $a\leqslant x<b$.
- $p(x)=\frac{1}{b-a}1_{\{a\leqslant x\leqslant b\}}$.
- $p(x)=\frac{1}{b-a}$, 其中$a\leqslant x\leqslant b$.
- 模型: 几何概型.

2. 指数分布, $X \sim \mathrm{Exp}(\lambda)$ (参数$\lambda > 0$):

$$p(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

- 模型：例2.3. X = 第一个粒子的放射时刻. 等待时间、寿命.
- 第一个粒子在第Y 个微观时间放出, 则

$$Y \sim G(p), \text{其中} p = \lambda \times \frac{1}{n}.$$



- $X \approx \frac{Y}{n}$, 于是,

$$P(X > t) \approx P(Y > nt) \approx (1-p)^{nt} = (1 - \frac{\lambda}{n})^{nt} \approx e^{-\lambda t}.$$

- $P(X > t) = e^{-\lambda t} = \int_t^{\infty} \lambda e^{-\lambda x} dx.$
- 定理3.1.(无记忆性): $P(X - s > t | X > s) = e^{-\lambda t}, \forall t, s \geqslant 0.$

3. 正态分布, $X \sim N(\mu, \sigma^2)$ (参数$\mu \in \mathbb{R}$, $\sigma > 0$):

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

- 标准正态分布$N(0,1)$:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

- §1.7. 第二近似公式: $x_k = \frac{k-np}{\sqrt{npq}}$,

$$C_n^k p^k q^{n-k} \approx \frac{1}{\sqrt{npq}} \phi(x_k).$$

- $\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, 将积分的平方写为二重积分:

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\times\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy=\frac{1}{2\pi}\iint\limits_{\mathbb{R}^2}e^{-\frac{x^2+y^2}{2}}dxdy.$$

- 做极坐标变换:

$$x=r\cos\theta,\ y=r\sin\theta\ \Rightarrow\begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial y}{\partial r}\\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta}\end{vmatrix}=r.$$

- 因此,

$$\star\star=\frac{1}{2\pi}\int_0^{2\pi}\left(\int_0^{\infty}e^{-\frac{r^2}{2}}rdr\right)d\theta=\int_0^{\infty}e^{-R}dR=1.$$

- $p(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$: 令$y=\frac{x-\mu}{\sigma}$, 则

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\cdot\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy=1.$$

- 函数Φ:

$$\Phi(x) = \int_{-\infty}^{x} \phi(z)dz.$$

- $\Phi(-x) = 1 - \Phi(x)$.
- 定理3.2. 令$x^* = \frac{x-\mu}{\sigma}$, 则

$$P(a < X < b) = \int_a^b \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)dx = \Phi(b^*) - \Phi(a^*).$$

- 推论3.1. 查表得$\Phi(3) = 0.9987$, 因此

$$P(\mu - 3\sigma < X < \mu + 3\sigma) = \Phi(3) - \Phi(-3) = 0.9974.$$

4. 威布尔(Weibull)分布, $X \sim \mathrm{W}(m,\eta)$ (参数$m,\eta>0$):

$$p(x)=\frac{m}{\eta^m}x^{m-1}\exp\left\{-\left(\frac{x}{\eta}\right)^m\right\},\quad x>0.$$

- $\int_0^\infty my^{m-1}e^{-y^m}dy=\int_0^\infty e^{-z}dz=1.$
- $m>0$: 形状参数; $\eta>0$: 尺度参数.
- $m=1$ 时就是指数分布$\mathrm{Exp}(\frac{1}{\eta})$.
- 应用: 机电产品的寿命; 可靠性研究.

5. 伽玛分布, $X \sim \Gamma(\alpha, \beta)$ (参数$\alpha, \beta > 0$):

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0.$$

- $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$.
- $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$:

$$\int_0^\infty y^\alpha e^{-y} dy = -y^\alpha e^{-y}|_0^\infty + \int_0^\infty \alpha y^{\alpha-1} e^{-y} dy.$$

- $\Gamma(1) = 1$; $\Gamma(\frac{1}{2}) = \sqrt{\pi}$:

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty \frac{1}{\sqrt{y}} e^{-y} dy = \sqrt{2} \int_0^\infty e^{-\frac{x^2}{2}} dx = \sqrt{\pi}.$$

- $\alpha = 1$ 时就是指数分布Exp(β).

§2.4 随机变量的严格定义与分布函数

- 定义4.1. 假设$(\Omega, \mathcal{F}, P)$ 是概率空间, $X : \Omega \to \mathbb{R}$ 满足:

$$对任意 x \in \mathbb{R} 都有 \{X \leqslant x\} \in \mathcal{F},$$

则称X 是一个随机变量.

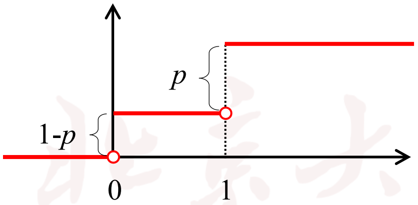
- 定义4.2. 令$F(x) = P(X \leqslant x)$, $x \in \mathbb{R}$. 称F 为随机变量X 的分布函数, 也记为F_X.
- 定理4.2. $F = F_X$ 的三条性质:

(1) 单调性: 若$x \leqslant y$, 则$F(x) \leqslant F(y)$.

(2) 规范性: $\lim_{x\to-\infty} F(x) = 0$; $\lim_{x\to\infty} F(x) = 1$.

(3) 右连续性: $\lim_{y\to x+} F(y) = F(x)$.

- 离散型: $P(X=x_i)=p_i$. x_i 为F_X 的跳点, p_i 为跳跃幅度.



- 连续型: $F_X(x)=\int_{-\infty}^{x}p(z)dz$, 且

$$p(x)=F_X'(x).$$

反过来, 若F_X "几乎"连续可导, 则为连续型(定理4.3, 4.4).

- 尾分布函数: $G(x) = P(X > x) = 1 - F(x)$.
  连续型: $p(x) = -G'(x)$.
- 例. $X \sim \mathrm{Exp}(\lambda)$.

$$G(x) = e^{-\lambda x}, \quad \forall x > 0,$$

$\Rightarrow G'(x) = -\lambda G(x)$. $\quad \lambda$: 速率.

- 由$F_X(x)$ 可求出$P(X \in B), \forall B$.
- 若$F_X = F_Y$, 则称X 与Y 同分布, 记为$X \overset{d}{=} Y$.
- $X = Y$, 即$P(X = Y) = 1$, 则$F_X = F_Y$. 反之不然.

§2.5 随机变量的函数

- 函数$f:\mathbb{R}\to\mathbb{R}$, $x\mapsto y=f(x)$. (定理5.1. 要求是Borel函数). 随机变量X 的函数指: 一个新的随机变量

$$Y=f(X):\omega\mapsto f(X(\omega)).$$

- 目标: 求Y 的分布列或密度函数.
- 若$X\overset{d}{=}\tilde{X}$, 则$f(X)\overset{d}{=}f(\tilde{X})$.
- 例. 假设X 的分布列为$P(X=x_i)=p_i,\forall i$. 则Y 也是离散型, 将其可能取值记为$y_j,\forall j$.

$$P(Y=y_j)=\sum_{i:f(x_i)=y_j}p_i.$$

例5.1. $X \sim N(\mu, \sigma^2)$, 求$Y = \frac{X-\mu}{\sigma}$ 的分布.

- 分布函数法. 第一步, 将$Y \leqslant y$ 改写为$X \leqslant \mu + \sigma y$. 从而,

$$F_Y(y) = P(Y \leqslant y) = P(X \leqslant \mu + \sigma y).$$

- 第二步, 代入X 的密度, 做变量替换:

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{\mu+\sigma y} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \\ &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt. \end{aligned}$$

- 第三步, 求导, 得到$p_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$, 即$Y \sim N(0,1)$.
- 一般地, $Z := a + bX \sim N(a + b\mu, b^2\sigma^2)$, 其中$b \neq 0$.

定理5.2. 由分布函数法推导密度变换公式.

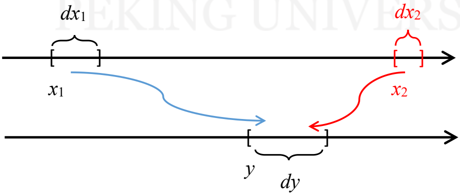
- 一对一情形, 假设f 连续可导, $f'(x) > 0, \forall x$.
- 反函数: $f(x) = y, x = g(y)$.
- $F_Y(y) = P(Y \leqslant y) = P(X \leqslant x) = \int_{-\infty}^{x} p_X(t)dt$.
- $p_Y(y) = F'_Y(y) = p_X(x)\frac{dx}{dy}$.
- 即$p_X(x)dx = p_Y(y)dy$.

- 一对一情形：定理5.2. 假设f 连续可导，$f'(x)>0$, $\forall x$, 则

$$p_X(x)dx = p_Y(y)dy \to p_Y(y) = p_X(x)\frac{1}{f'(x)} = p_X(g(y))g'(y).$$

- 多对一情形*：f 为分段的一对一情形的函数. 例如，f 是多项式函数，都有

$$p_Y(y) = \sum_{x_i: f(x_i)=y} p_X(x_i)\frac{1}{|f'(x_i)|}.$$



例5.2. 假设$X \sim N(0,1)$, $Y = X^2$. 求Y 的密度函数.

- 方法一、分布函数法: 对任意$y > 0$,

$$F_Y(y) = P(-\sqrt{y} \leqslant X \leqslant \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$\Rightarrow p_Y(y) = p_X(\sqrt{y})\frac{d\sqrt{y}}{dy} - p_X(-\sqrt{y})\frac{d(-\sqrt{y})}{dy} = \frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}.$$

- 方法二、用多对一公式:

  第一步, 确定每个$y > 0$ 的原像点: $x_1 = \sqrt{y}$, $x_2 = -\sqrt{y}$.

- 第二步, 求出每个x_i 的贡献: $p_X(x_i)\left|\frac{dx_i}{dy}\right| = \frac{1}{2\sqrt{2\pi y}}e^{-\frac{y}{2}}$.
- 第三步, 对i 求和:

$$p_Y(y) = \sum_{i=1}^{2} p_X(x_i)\left|\frac{dx_i}{dy}\right| = \frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}, \quad \text{其中} y > 0.$$

- 注: $X^2 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$.

反例. 假设$X \sim N(0,1)$, $Y = f(X)$, 其中,

$$f(x) = \begin{cases} x^2, & 若\ |x| > 1; \\ 1, & 若\ |x| \leqslant 1. \end{cases}$$

- Y 不是连续型:

$$P(Y = 1) = P(|X| \leqslant 1) > 0.$$

- f 在$(-1,1)$ 上恒有$f'(x) = 0$.

例5.4. 已知$Y = \ln X \sim N(\mu, \sigma^2)$. 求$X$ 的分布.

- 称X 服从对数正态分布.
- $X = e^Y > 0$. 因此, $\forall x > 0$, 有

$$\begin{aligned} G_X(x) &= P(X > x) = P(Y > \ln x) \\ &= \int_{\ln x}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\ &= \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\ln u-\mu)^2}{2\sigma^2}} \frac{1}{u} du. \end{aligned}$$

- $p_X(x) = -G_X'(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \ x > 0.$

- 定义5.1. 设F 是分布函数(满足单调、规范、右连续). 令

$$F^{-1}(p) := \inf\{x : F(x) \geqslant p\}, \quad \forall p \in (0,1).$$

称F^{-1} 为F 的广义反函数.

- 定理5.3. 假设F 是分布函数.

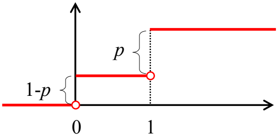
若$U \sim U(0,1)$, 则$X = F^{-1}(U)$ 满足$F_X = F$.

- 证明: $F^{-1}(p) \leqslant x$ 当且仅当$p \leqslant F(x)$. 于是,

$$P(X \leqslant x) = P(F^{-1}(U) \leqslant x) = P(U \leqslant F(x)) = F(x).$$

- 例5.3. 假设F_X 连续, 则

$$Y = F_X(X) \sim U(0,1).$$



- F_X 不连续时的反例, $X \sim B(1,p)$.

§2.6 随机变量的数学期望

期望(expectation)的含义: 均值(mean).

- X 的大量独立观测值(记为$a_1, a_2, \cdots, a_n$) 的算术平均:
$$\bar{a} = \frac{1}{n}(a_1 + \cdots + a_n).$$
- X 的所有可能值的加权平均(总和).

  例, $P(X = x_k) = p_k$, $k = 1, \cdots, m$.

  记$n_k = \{m : 1 \leqslant m \leqslant n, a_m = x_k\}$. 那么, 根据概率的频率含义, $\frac{n_k}{n} \approx p_k$, 于是

$$\bar{a} = \frac{1}{n}\sum_{k=1}^{K} n_k x_k \approx \sum_{k=1}^{K} x_k p_k.$$

1. 离散型随机变量的期望

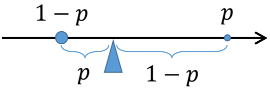
- 定义6.1. 假设X 是离散型, 分布列为

$$P(X = x_k) = p_k, \quad k = 1, \cdots, n \ \text{或}\ k = 1, 2, \cdots .$$

如果$\sum_k |x_k| p_k < \infty$, 那么, 称$X$ 的期望存在, 称$\sum_k x_k p_k$ 为X 的数学期望, 记为EX.

- EX 是重心.

例, (1) 伯努利分布, $P(X = 1) = p$, $P(X = 0) = 1 - p$. 则, $EX = p$.



(3) 泊松分布.

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}=:p_k,\quad k=0,1,2,\cdots.$$

- $\forall k\geqslant 1$, x_k,

$$kp_k=k\frac{\lambda^k}{k!}e^{-\lambda}=\lambda\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}=\lambda p_{k-1}.$$

- 因此,

$$EX=\sum_{k=0}^{\infty}kp_k=\sum_{k=1}^{\infty}\lambda p_{k-1}=\lambda\sum_{\ell=0}^{\infty}p_\ell=\lambda.$$

(2) 二项分布.

$$P(X=k)=C_n^k p^k q^{n-k} =: b(n;k),\ \ k=0,1,\cdots,n,(q=1-p).$$

- $\forall 1\leqslant k\leqslant n$,

$$\begin{aligned} k\cdot b(n;k) =& k\cdot\frac{n!}{k!(n-k)!}p^k q^{n-k}=\frac{n!}{(k-1)!(n-k)!}p^k q^{n-k}\\ =&\frac{n\cdot(n-1)!}{(k-1)!(n-k)!}p\cdot p^{k-1}q^{n-k}=np\cdot b(n-1,k-1). \end{aligned}$$

- 因此,

$$\begin{aligned} EX=&\sum_{k=0}^{n}k\cdot b(n;k)=\sum_{k=1}^{n}np\cdot b(n-1;k-1)\\ =&np\sum_{\ell=0}^{n-1}b(n-1,\ell)=np. \end{aligned}$$

(7) 超几何分布.

$$P(X=k)=\frac{C_D^k C_{N-D}^{n-k}}{C_N^n},\quad k=0,1,\cdots,n.$$

- 记$h(N,D,n;k)=A_1\cdot A_2\cdot A_3=$

$$\frac{D!}{k!(D-k)!}\cdot\frac{(N-D)!}{(n-k)!\big(N-D-(n-k)\big)!}\cdot\frac{n!(N-n)!}{N!}.$$

- 记$x'=x-1$. 则, $\forall 1\leqslant k\leqslant n$,

$$k\cdot A_1=\frac{D!}{(k-1)!(D-k)!}=D\times\frac{D'!}{k'!\big(D'-k'\big)!}.$$

- 进一步,

$$A_2=\frac{(N'-D')!}{(n'-k')!\big(N'-D'-(n'-k')\big)!},$$

$$A_3=\frac{n\cdot n'!(N'-n')!}{N\cdot N'!}=\frac{n}{N}\times\frac{n'!(N'-n')!}{N'!}.$$

- 记$x' = x - 1$. 则$\forall 1 \leqslant k \leqslant n$,

$$k \cdot h(N, D, n; k) = \frac{nD}{N} \times h(N', D', n'; k').$$

- 因此,

$$EX = \sum_{k=1}^{n} k \cdot h(N, D, n; k) = \frac{nD}{N} \sum_{k'=0}^{n'} h(N', D', n'; k') = \frac{nD}{N}.$$

- $D = 1$ 时, 退化为伯努利分布, $EX = p = \frac{D}{N}$.
- $D \geqslant 2$ 时, 不放回抽样, 仍有$EX = np$.

(4) 几何分布.

$$P(X=k)=q^{k-1}p=:p_k,\quad k=1,2,\cdots,(q=1-p).$$

- 直接计算:

$$EX=\sum_{k=1}^{\infty}kp_k=\sum_{k=1}^{\infty}\sum_{\ell=1}^{k}p_k=\sum_{\ell=1}^{\infty}\sum_{k=\ell}^{\infty}p_k$$
$$=\sum_{\ell=1}^{\infty}p\cdot\frac{q^{\ell-1}}{1-q}=\sum_{m=0}^{\infty}q^m=\frac{1}{1-q}=\frac{1}{p}.$$

- 习题二、18. 若X 取非负整数, 则$EX=\sum_{\ell=1}^{\infty}P(X\geqslant\ell)$.
- 证: $\sum_{k=\ell}^{\infty}p_k=P(X\geqslant\ell)$.

2. 一般随机变量的期望

- X 为任意随机变量. 做如下近似: $\forall n \in \mathbb{Z}$,

$$\text{当 } n\varepsilon \leqslant X < (n+1)\varepsilon \text{ 时, 令 } Y = n\varepsilon.$$

- 直观: $Y \leqslant X < Y + \varepsilon$, 因此$EY \leqslant EX < EY + \varepsilon$.
- 定义6.2. 若EY 存在且当$\varepsilon \to 0$ 时有极限, 则称X 的期望存在, 且称该极限为X 的期望, 记为EX.
- 对离散型随机变量, 定义6.1与定义6.2一致.
- 定理6.1. 对连续型随机变量, 若$\int_{-\infty}^{\infty} |x|p(x)dx < \infty$, 则

$$EX = \int_{-\infty}^{\infty} xp(x)dx.$$

(2) 指数分布.

$$p(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

- $\int_0^\infty x \cdot \lambda e^{-\lambda x} dx = -\int_0^\infty x de^{-\lambda x} = \int_0^\infty e^{-\lambda x} dx = \frac{1}{\lambda}.$
- 一般地, 若X 为连续型, 且$X \geqslant 0$. 令

$$G(x) = P(X > x) = \int_x^\infty p(y)dy,$$

则$G'(x) = -p(x)$. 于是,

$$\int_0^\infty xp(x)dx = -\int_0^\infty xdG(x) = \int_0^\infty G(x)dx.$$

(3) 正态分布.

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

- $X \sim N(0,1)$:

$$EX = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0.$$

- 同理, $X \sim N(\mu, \sigma^2)$, 则$p(\mu + x) = p(\mu - x)$, 因此$EX = \mu$.
- 例, 柯西分布,

$$p(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

但是, $\int_{-\infty}^{\infty} |x| p(x) dx = \infty$. 因此, EX 不存在!

(4) 伽玛分布.

$$p(x)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x},\quad x>0.$$

- $\forall x>0$,

$$xp(x)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha}e^{-\beta x}=\frac{\Gamma(\alpha+1)}{\beta\Gamma(\alpha)}\cdot\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}x^{\alpha}e^{-\beta x}=\frac{\alpha}{\beta}\cdot\hat{p}(x).$$

- 因此,

$$EX=\int_0^{\infty}xp(x)dx=\frac{\alpha}{\beta}\int_0^{\infty}\hat{p}(x)dx=\frac{\alpha}{\beta}.$$

3. 期望的性质

- 定理6.2. (1) 若$X \equiv a$, 则$EX = a$;
- 定理6.2. (2) 若$X \geqslant 0$, 且EX 存在, 则$EX \geqslant 0$;
- 定理6.2. (3)(或, 推论6.1). 若$F_X = F_Y$ (或, 若$X = Y$), 且EX 存在, 则EY 存在, 且$EX = EY$.
- 定理6.3. (1) & (2), 线性: 假设EX, EY 存在. 则,

$$E(aX) = aEX, \quad E(X+Y) = EX + EY.$$

- 定理6.3. (3), 单调性: 假设⋆⋆. 又若$X \geqslant Y$, 则$EX \geqslant EY$.

- 推论6.2. (1) 线性: 假设EX, EY 存在. 则,

$$E(aX+bY)=aEX+bEY.$$

- 推论6.2. (2) 和的期望: 假设$EX_1,\cdots,EX_n$ 都存在, $\eta=X_1+\cdots+X_n$. 则$E\eta$ 存在, 且

$$E\eta=EX_1+\cdots+EX_n.$$

- 例. 超几何分布$\eta\sim H(N,D,n)$.
  若第i 个产品是次品, 则令$X_i=1$; 否则, 令$X_i=0$. 则,

$$\eta=X_1+\cdots+X_n\Rightarrow E\eta=np.$$

- 定理6.4. (马尔可夫不等式). 设$X \geqslant 0$, 且EX 存在. 则对任意$C > 0$, 有

$$P(X \geqslant C) \leqslant \frac{1}{C} EX.$$

- 证: 令$A = \{X \geqslant C\}$. 则$1_A \leqslant \frac{X}{C}$. 于是,

$$P(A) = E1_A \leqslant E\frac{X}{C} = \frac{1}{C} EX.$$

- 例, 若$X \geqslant 0$, 且$EX = 0$, 则

$$P\left(X \geqslant \frac{1}{n}\right) \leqslant nEX = 0$$

$$\Rightarrow P(X > 0) = \lim_{n\to\infty} P\left(X \geqslant \frac{1}{n}\right) = 0.$$

- 定理6.5. X 是离散型，或连续型，且下面的级数或积分绝对收敛，则

$$Ef(X)=\sum_k f(x_k)p_k, \quad \text{或 } Ef(X)=\int_{-\infty}^{\infty} f(x)p(x)dx.$$

- 例6.1. 设$X \sim U(0,2\pi)$，求$E\sin X$.
- 用公式:

$$E\sin X=\int_{-\infty}^{\infty}\sin x\cdot p(x)dx=\frac{1}{2\pi}\int_0^{2\pi}\sin x dx=0.$$

- 定义7.1 & 7.2.. 假设EX 存在, 且$E(X-EX)^2$ 也存在. 则称$E(X-EX)^2$ 为X 的方差, 记为$\mathrm{var}(X)$ 或$D(X)$.
  称$\sqrt{\mathrm{var}(X)}$ 为标准差.
- 记$\mu=\mu_X=EX$, $\sigma=\sigma_X=\sqrt{D(X)}$, $\sigma^2=\sigma_X^2=D(X)$.
- 定理7.1 (切比雪夫不等式) 假设$\mathrm{var}(X)<\infty$, 则$\forall\varepsilon>0$,

$$P(|X-EX|\geqslant\varepsilon)\leqslant\frac{1}{\varepsilon^2}\mathrm{var}(X).$$

- 证: $\{|X-EX|\geqslant\varepsilon\}=\{(X-EX)^2\geqslant\varepsilon^2\}$,
  对$Y=(X-EX)^2$ 用马尔可夫不等式.
- 推论7.1. 若$\mathrm{var}(X)=0$, 则X 退化.
- 证: $Y\geqslant 0$ 且$EY=0$, 故$Y\equiv 0$, 即$X\equiv c=EX$.

- 定理7.2. $\mathrm{var}(X) = EX^2 - (EX)^2$.
- 证: $(X-\mu)^2 = X^2 - 2\mu X + \mu^2$, 故

$$\mathrm{var}(X) = EX^2 - 2\mu EX + \mu^2 = EX^2 - \mu^2.$$

- 具体地, 离散型或连续型的公式如下:

$$\mathrm{var}(X) = \sum_k x_k^2 p_k - (EX)^2,$$

$$\mathrm{var}(X) = \int_{-\infty}^{\infty} x^2 p(x)dx - (EX)^2.$$

- X 的线性变换的方差:

$$\mathrm{var}(aX+b) = a^2\mathrm{var}(X).$$

(3) 泊松分布.

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}, \quad k=0,1,2,\cdots.$$

- $EX=\lambda$, 且$\forall k\geqslant 1$, $kp_k=\lambda p_{k-1}$. 因此, $\forall k\geqslant 2$,

$$k^2p_k=\lambda kp_{k-1}=\lambda p_{k-1}+\lambda(k-1)p_{k-1}=\lambda p_{k-1}+\lambda^2p_{k-2}.$$

- 或者, $\forall k\geqslant 2$,

$$k(k-1)p_k=\lambda(k-1)p_{k-1}=\lambda^2p_{k-2}.$$

- 于是, $EX(X-1)=\lambda^2$, 从而

$$\mathrm{var}(X)=EX^2-(EX)^2=EX(X-1)+EX-(EX)^2=\lambda.$$

(2) 二项分布.

$$P(X=k)=C_n^k p^k q^{n-k} =: b(n;k),\ \ k=0,1,\cdots,n,(q=1-p).$$

- $EX=np$, 且$\forall 1\leqslant k\leqslant n$,

$$k\cdot b(n;k)=np\cdot b(n-1,k-1).$$

- $\forall 2\leqslant k\leqslant n$,

$$\begin{aligned} k(k-1)\cdot b(n;k) &= np\cdot(k-1)\cdot b(n-1,k-1) \\ &= np\cdot(n-1)p\cdot b(n-2,k-2) \end{aligned}$$

- 于是, $EX(X-1)=np(n-1)p=(np)^2-np^2$, 从而

$$D(X)=EX^2-(EX)^2=EX(X-1)+EX-(EX)^2=npq.$$

(6) 正态分布.

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

- 若$\mu = EX = 0$, $\sigma^2 = 1$, 则,

$$\begin{aligned}\mathrm{var}(X) =& EX^2 = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ =& -\frac{2}{\sqrt{2\pi}} \int_0^{\infty} x d e^{-\frac{x^2}{2}} = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{x^2}{2}} dx = 1.\end{aligned}$$

- 一般情形, $Y = \frac{X-\mu}{\sigma} \sim N(0,1)$. 则

$$\mathrm{var}(X) = \mathrm{var}(\mu + \sigma Y) = \sigma^2 \mathrm{var}(Y) = \sigma^2.$$

- 一般地, 若X 的方差存在, 且$\mathrm{var}(X) > 0$, 则

$$X^* = \frac{X - EX}{\sqrt{\mathrm{var}(X)}}$$

满足$EX^* = 0$, $\mathrm{var}(X^*) = 1$. 称X^* 为X 的标准化.

- 定义7.3. k 阶(原点)矩指EX^k.

定义7.4. k 阶中心矩指$E(X - EX)^k$.

- 定义7.5. 若

$$P(X < a) \leqslant p \leqslant P(X \leqslant a),$$

则称a 为X 的一个p 分位数.

$p = 0.5$ 时, 也称a 为一个中位数.