660概率论第二，三章选择题思维导图

分布函数及密度函数的计算

440，442，510

分布函数法 — $F_Y(y)=P\{Y\leqslant y\}=P\{g(X)\leqslant y\}=\int_{g(x)\leqslant y} f_X(x)\mathrm{d}x$

公式法 — $f_Y(y)=\begin{cases} f_X[h(y)]\cdot|h'(y)|, & \alpha<y<\beta \\ 0, & 其他 \end{cases}$

设 X 为离散型随机变量，其概率分布为 $p_i=P\{X=x_i\}(i=1,2,\cdots)$，则 X 的函数 $Y=g(X)$ 也是离散型随机变量，其概率分布为 $P\{Y=g(x_i)\}=p_i$，即

$$Y\sim\begin{bmatrix} g(x_1) & g(x_2) & \cdots \\ p_1 & p_2 & \cdots \end{bmatrix}.$$

如果有若干个 $g(x_k)$ 相同，则合并诸项为一项 $g(x_k)$，并将相应概率相加作为 Y 取 $g(x_k)$ 值的概率.

常见连续型分布

498，505

① 均匀分布 $U(a,b)$.

如果随机变量 X 的概率密度或分布函数分别为

$$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b, \\ 0, & 其他,\end{cases} \quad F(x)=\begin{cases} 0, & x<a, \\ \frac{x-a}{b-a}, & a\leqslant x<b, \\ 1, & x\geqslant b,\end{cases}$$

435，437，443，508

② 指数分布 $E(\lambda)$.

如果 X 的概率密度或分布函数分别为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0, \\ 0, & 其他\end{cases}(\lambda>0), F(x)=\begin{cases}1-e^{-\lambda x}, & x\geqslant 0, \\ 0, & x<0\end{cases}(\lambda>0),$$

则称 X 服从参数为 λ 的**指数分布**，记为 $X\sim E(\lambda)$.（如图 2-5,2-6）

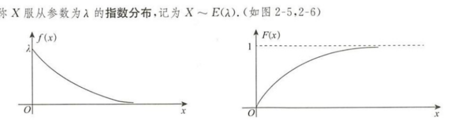

图 2-5　　图 2-6

注意：1° $P\{X\geqslant t+s \mid X\geqslant t\}=P\{X\geqslant s\}$ 无记忆性.

2° $F(x)=\begin{cases}1-e^{-\lambda x}, x\geqslant 0, \\ 0, \quad x<0\end{cases}(\lambda>0)$（记，易考）.

③ 正态分布.

若 $X\sim f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty<x<+\infty$，其中 $-\infty<\mu<+\infty, \sigma>0$，

则称 X 服从参数为 (μ,σ^2) 的正态分布，记为 $X\sim N(\mu,\sigma^2)$.

注意：1° $\mu=0,\sigma=1$ 时的正态分布 $N(0,1)$ 为标准正态分布，

$$X\sim\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t,$$

则 $X\sim N(0,1)$.

444，509，516

3° 若 $X\sim N(\mu,\sigma^2)$，则

$\frac{X-\mu}{\sigma}\sim N(0,1)$；

$F(x)=P\{X\leqslant x\}=\Phi\left(\frac{x-\mu}{\sigma}\right)$；

$P\{a\leqslant X\leqslant b\}=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$；

$P\{\mu-\sigma\leqslant X\leqslant\mu+\sigma\}=2\Phi(1)-1$；

$P\{\mu-k\sigma\leqslant X\leqslant\mu+k\sigma\}=2\Phi(k)-1(k>0)$.

4° 若 $X\sim N(0,1)$，则

$\Phi(-x)=1-\Phi(x)$；

$P\{|X|\leqslant a\}=2\Phi(a)-1(a>0)$；

$P\{|X|>a\}=2[1-\Phi(a)](a>0)$.

多维随机变量及其分布

对任意的实数 x,y，称二元函数

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}$$

为二维随机变量 (X,Y) 的**联合分布函数**，记为 $(X,Y)\sim F(x,y)$. $F(x,y)$ 是事件 $A=\{X\leqslant x\}$ 与 $B=\{Y\leqslant y\}$ 同时发生的概率.

(1) $F(x,y)$ 是联合分布函数的充要条件.

① **单调性**　$F(x,y)$ 是 x,y 的单调不减函数：

当 $x_1<x_2$ 时，$F(x_1,y)\leqslant F(x_2,y)$；当 $y_1<y_2$ 时，$F(x,y_1)\leqslant F(x,y_2)$.

② **右连续性**　$F(x,y)$ 是 x,y 的右连续函数：

$$\lim_{x\to x_0^+}F(x,y)=F(x_0+0,y)=F(x_0,y);$$
$$\lim_{y\to y_0^+}F(x,y)=F(x,y_0+0)=F(x,y_0).$$

③ **有界性**　$F(-\infty,y)=F(x,-\infty)=F(-\infty,-\infty)=0, F(+\infty,+\infty)=1$.

④ **非负性**　对任意 $x_1<x_2,y_1<y_2$ 有

$$P\{x_1<X\leqslant x_2,y_1<Y\leqslant y_2\}=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\geqslant 0.$$

$f(x,y)$ 是联合概率密度的充要条件.

如果二维随机变量 (X,Y) 的联合分布函数 $F(x,y)$ 可以表示为

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v,(x,y)\in\mathbf{R}^2,$$

其中 $f(x,y)$ 是非负可积函数，则称 (X,Y) 为**二维连续型随机变量**，称 $f(x,y)$ 为 (X,Y) 的**概率密度**，记为 $(X,Y)\sim f(x,y)$.

$f(x,y)$ 是联合概率密度的充要条件为 $f(x,y)\geqslant 0$ 且 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1$.

(4) 反问题（重点）.

用 $\begin{cases}F(-\infty,y)=0,F(x,-\infty)=0, \\ F(-\infty,-\infty)=0,F(+\infty,+\infty)=1, \\ \sum_j\sum_i p_{ij}=1,\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1\end{cases}$ 建方程，求参数.

526

求联合分布

离散型F（x,y）=P{X≤x，Y≤y}=pij

连续型

b. $(X,Y)\sim f(x,y)$，则

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}=\int_{-\infty}^{x}\mathrm{d}u\int_{-\infty}^{y}f(u,v)\mathrm{d}v.$$

a. 二维均匀分布.

称 (X,Y) 在平面有界区域 D 上服从**均匀分布**，如果 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\frac{1}{S_D}, & (x,y)\in D, \\ 0, & 其他,\end{cases}$$

其中 S_D 为区域 D 的面积.

b. 二维正态分布.

如果 (X,Y) 的概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\},$$

其中 $\mu_1\in\mathbf{R},\mu_2\in\mathbf{R},\sigma_1>0,\sigma_2>0,-1<\rho<1$，则称 (X,Y) 服从参数为 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$ 的**二维正态分布**，记为 $(X,Y)\sim N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$.

452，453，520

(2) 求边缘分布.

① 求 $F_X(x),F_Y(y)$.

$$F_X(x)=F(x,+\infty),F_Y(y)=F(+\infty,y).$$

② 求 $p_{i\cdot},p_{\cdot j}$.

$$p_{i\cdot}=\sum_j p_{ij},p_{\cdot j}=\sum_i p_{ij}.$$

③ 求 $f_X(x),f_Y(y)$.

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y,f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x.$$

521

(3) 求条件分布.

① 求 $P\{Y=y_j \mid X=x_i\},P\{X=x_i \mid Y=y_j\}$.

$$P\{Y=y_j \mid X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_{i\cdot}},$$
$$P\{X=x_i \mid Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}}.$$

② 求 $f_{Y|X}(y \mid x),f_{X|Y}(x \mid y)$.

$$f_{Y|X}(y \mid x)=\frac{f(x,y)}{f_X(x)},f_{X|Y}(x \mid y)=\frac{f(x,y)}{f_Y(y)}.$$

446，517

随机变量及其分布函数

随机变量

设随机试验 E 的样本空间为 $\Omega=\{\omega\}$，如果对每一个 $\omega\in\Omega$，都有唯一的实数 $X(\omega)$ 与之对应，并且对任意实数 x，$\{\omega \mid X(\omega)\leqslant x,\omega\in\Omega\}$ 是随机事件，则称定义在 Ω 上的实值单值函数 $X(\omega)$ 为**随机变量**，简记为随机变量 X.

分布函数

(2) 分布函数.

设 X 是随机变量，x 是任意实数，称函数 $F(x)=P\{X\leqslant x\}(x\in\mathbf{R})$ 为随机变量 X 的分布函数，或称 X 服从 $F(x)$ 分布，记为 $X\sim F(x)$.

离散型随机变量

(1) 离散型随机变量及其概率分布.

如果随机变量 X 只可能取有限个或可列个值 $x_1,x_2,\cdots$，则称 X 为**离散型随机变量**，称

$$p_i=P\{X=x_i\},i=1,2,\cdots$$

为 X 的**分布列**、**分布律**或**概率分布**，记为 $X\sim p_i$，概率分布常常用表格形式或矩阵形式表示，即

X	x_1	x_2	$\cdots$
P	p_1	p_2	$\cdots$

或 $X\sim\begin{bmatrix}x_1 & x_2 & \cdots \\ p_1 & p_2 & \cdots\end{bmatrix}$.

519

连续型随机变量

(2) 连续型随机变量及其概率密度.

如果随机变量 X 的分布函数可以表示为

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t(x\in\mathbf{R}),$$

其中 $f(x)$ 是非负可积函数，则称 X 为**连续型随机变量**，称 $f(x)$ 为 X 的**概率密度函数**，简称**概率密度**，记为 $X\sim f(x)$.

分布函数和密度函数的性质

(1) $F(x)$ 是分布函数 $\Leftrightarrow F(x)$ 是 x 的单调不减、右连续函数，且 $F(-\infty)=0,F(+\infty)=1$.

(2) $\{p_i\}$ 是概率分布 $\Leftrightarrow p_i\geqslant 0$，且 $\sum_i p_i=1$.

(3) $f(x)$ 是概率密度 $\Leftrightarrow f(x)\geqslant 0$，且 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$.

441，497，499，500，501，502，503，504，506，507

常见离散型分布

(2) 常见以下 5 种离散型分布.

① 0－1 分布.

$X\sim B(1,p)$，X（伯努利计数变量）$\sim\begin{pmatrix}1 & 0 \\ p & 1-p\end{pmatrix}$.

② 二项分布.

$X\sim B(n,p)\begin{cases}a.\ n 次试验相互独立; \\ b.\ P(A)=p; \\ c.\ 只有 A,\bar{A} 两种结果.\end{cases}$

记 X 为 A 发生的次数，则

$$P\{X=k\}=C_n^k\cdot p^k(1-p)^{n-k},k=0,1,2,\cdots,n.$$

439，518

③ 几何分布.

$X\sim G(p)$ 首中即停止（等待型分布），记 X 为试验次数，则

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

④ 超几何分布.

N 件产品中有 M 件正品，无放回取 n 次，则取到 k 个正品的概率

$$P\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k 为整数, \max\{0,n-N+M\}\leqslant k\leqslant\min\{n,M\}.$$

⑤ 泊松分布.

某单位时间段，某场合下，源源不断的随机质点流的个数，也常用于描述稀有事件的概率.

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}(k=0,1,\cdots;\lambda>0), \lambda 表示强度(EX=\lambda).$$

泊松定理　若 $X\sim B(n,p)$，当 n 很大，p 很小，$\lambda=np$ 适中时，二项分布可用泊松分布近似表示，

即 $$C_n^k p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda}.$$

分布的常见运算公式

① $P\{X\leqslant a\}=F(a)$；

② $P\{X<a\}=F(a-0)$；

③ $P\{X=a\}=P\{X\leqslant a\}-P\{X<a\}=F(a)-F(a-0)$；

④ $P\{a<X<b\}=P\{X<b\}-P\{X\leqslant a\}=F(b-0)-F(a)$；

⑤ $P\{a\leqslant X\leqslant b\}=P\{X\leqslant b\}-P\{X<a\}=F(b)-F(a-0)$.

522

用分布求概率

用分布求概率及反问题.

① $(X,Y)\sim p_{ij}$，则 $P\{(X,Y)\in D\}=\sum_{(x_i,y_j)\in D}p_{ij}$.

② $(X,Y)\sim f(x,y)$，则 $P\{(X,Y)\in D\}=\iint_D f(x,y)\mathrm{d}x\mathrm{d}y$.

③ (X,Y) 为混合型，则用全概率公式.

④ 反问题：已知概率反求参数.

445，447，448，449，454，511，512，513，514，515，523，524，525，527，528