3 多维随机变量及其分布

2020年7月31日 星期五 下午2:38

多维随机变量及其分布

- 多维随机变量及其分布函数
  - 多维随机变量的概念：如果 $X_1, X_2, \cdots, X_n$ 是定义在同一个样本空间 Ω 上的 n 个随机变量，则称 $(X_1, X_2, \cdots, X_n)$ 为 n 维随机变量或 n 维随机向量，$X_i\ (i=1,2,\cdots,n)$ 为第 i 个分量。
    当 $n=2$ 时，记 (X,Y) 为二维随机变量（或二维随机向量）
  - 联合分布函数
    - 概念：对任意的 n 个实数 $x_1, x_2, \cdots, x_n$，称 n 元函数 $F(x_1,x_2,\cdots,x_n) = P\{X_1 \le x_1, X_2 \le x_2, \cdots$ 为 n 维随机变量 $(X_1,X_2,\cdots,X_n)$ 的联合分布函数。
      当 $n=2$ 时，则对任意的实数 x,y，称二元函数 $F(x,y) = P\{X\le x, Y\le y\}$ 为二维 (X,Y) 的联合分布函数，简称分布函数，记为 $(X,Y) \sim F(x,y)$
    - 性质
      - 单调性：$F(x,y)$ 是 x,y 的单调不减函数 $\begin{cases} 固定y,\ x_1<x_2 \Rightarrow F(x_1,y) \le F(x_2,y) \\ 固定x,\ y_1<y_2 \Rightarrow F(x,y_1) \le F(x,y_2) \end{cases}$
      - 右连续性：$F(x,y)$ 是 x,y 的右连续函数 $\begin{cases} \lim_{x\to x_0^+} F(x,y) = F(x_0+0, y) = F(x_0,y) \\ \lim_{y\to y_0^+} F(x,y) = F(x, y_0+0) = F(x,y_0) \end{cases}$
      - 有界性：$F(-\infty, y) = F(x,-\infty) = F(-\infty,-\infty) = 0$；$F(+\infty,+\infty) = 1$
      - 非负性：对任意 $x_1<x_2$，$y_1<y_2$ 有 $P\{x_1<X\le x_2, y_1<Y\le y_2\} = F(x_2,y_2) - F(x$
  - 边缘分布函数：设二维随机变量 (X,Y) 的联合分布函数为 $F(x,y)$，随机变量 X 与 Y 的分布函数 $F_X(x)$ 与 $F_Y(y)$ 分 (X,Y) 关于 X 和关于 Y 的边缘分布函数，由概率性质得
    $F_X(x) = P\{X\le x\} = P\{X\le x, Y<+\infty\} = \lim_{y\to+\infty} F(x,y) = F(x,+\infty)$
- 二维离散型随机变量
  - 二维离散型随机变量及其概率分布
    - 概念：如果二维随机变量 (X,Y) 的可能取值为有限个或可数无穷个 $(x_i,y_j)\ (i=$ 二维离散型随机变量，称
      $p_{ij} = P\{X=x_i, Y=y_j\}$，$i,j=1,2,\cdots$
      为 (X,Y) 的概率分布或分布律，或随机变量 X 与 Y 的联合分布律，记为 $($
    - 性质：① $p_{ij} \ge 0$，$i,j=1,2,\cdots$；② $\sum_i\sum_j p_{ij} = 1$.（充分必要条件）
  - 联合分布函数：设 (X,Y) 的概率分布为 p_{ij}，$i,j=1,2,\cdots$，则 (X,Y) 的联合分布函数为 $F(x,y) = P\{X\le x, Y\le y\} = \sum$
    设 G 是平面上的某个区域，则 $P\{(X,Y)\in G\} = \sum_{(x_i,y_j)\in G} p_{ij}$
  - 边缘分布：X, Y 的边缘分布分别为 $\begin{cases} p_{i\cdot} = P\{X=x_i\} = \sum_{j=1}^{\infty} P\{X=x_i, Y=y_j\} = \sum_{j=1}^{\infty} p_{ij}\ (i=1,2,\cdots) \\ p_{\cdot j} = P\{Y=y_j\} = \sum_{i=1}^{\infty} P\{X=x_i, Y=y_j\} = \sum_{i=1}^{\infty} p_{ij}\ (j=1,2,\cdots) \end{cases}$
  - 条件分布：如果 $(X,Y) \sim p_{ij}\ (i,j=1,2,\cdots)$，对固定的 j，如果 $p_{\cdot j} = P\{Y=y_j\} > 0$，则称
    $$P\{X=x_i \mid Y=y_j\} = \frac{P\{X=x_i, Y=y_j\}}{P\{Y=y_j\}} = \frac{p_{ij}}{p_{\cdot j}}\quad (i=1,2,\cdots)$$
    为 X 在 "$Y=y_j$" 条件下的条件分布。

- 二维连续型随机变量
  - 二维连续性随机变量及其概率密度
    - 概念：如果二维随机变量(X,Y)的联合分布函数可表示为 $F(x,y)$ 其中 $f(x,y)$ 为非负可积函数，则称(X,Y)为二维连续型随
    - 性质：① $f(x,y)\geq 0$ ；② $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy=1$. ⇒ $f(x,y)$是概率
  - 联合分布函数与概率密度
    - ① $F(x,y)=P\{X\leq x, Y\leq y\}=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv$.
    - ② 设G为平面上的某个区域，则 $P\{(X,Y)\in G\}=\iint_G f(x,y)\,dx\,dy$
    - ③ 若$f(x,y)$在点(x,y)处连续，则 $\frac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y)$.
    - ④ 若$F(x,y)$连续且可导，则(X,Y)是连续型随机变量，则 $\frac{\partial^2 F(x,y)}{\partial x\partial y}$ 是它的概率密度
  - 边缘概率密度：设$(X,Y)\sim f(x,y)$，则X的边缘分布函数为 $F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty} f(u,v)\,dv\right]du$
    ∴ X是连续型随机变量，其概率密度 $f_X(x)=\int_{-\infty}^{+\infty} f(x,v)\,dv=\int_{-\infty}^{+\infty} f(x,y)\,dy$ ，称f_X(x
  - 条件概率密度：设$(X,Y)\sim f(x,y)$，边缘概率密度$f_X(x)>0$，则称 $f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}$ 为Y在"X=x"条件
    ⇒ $F_{Y|X}(y|x)=\int_{-\infty}^{y} f_{Y|X}(y|x)\,dy=\int_{-\infty}^{y}\frac{f(x,y)}{f_X(x)}\,dy$. Y在"X=x"条件下的条件分布函数.
  - 常见的二维分布
    - 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
      $$f(x,y)=\begin{cases}\frac{1}{S_D}, & (x,y)\in D\\ 0, & 其他\end{cases}$$
    - 二维正态分布
      - 概念：如果二维连续型随机变量(X,Y)的概率密度为
        $$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\right.}$$
        其中$\mu_1,\mu_2,\sigma_1>0,\sigma_2>0,-1<\rho<1$均为常数，则称(
        二维正态分布，记作 $(X,Y)\sim N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;$
      - 性质
        - ① $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，ρ为X与Y的相关系数，即 ρ
        - ② X,Y的条件分布都是正态分布
        - ③ $aX+bY$ ($a\neq 0$或$b\neq 0$)服从正态分布
        - ④ X与Y相互独立的充分必要条件是X与Y不相关，即$\rho=0$

- 随机变量的独立性
  - 概念：设二维随机变量(X,Y)的联合分布函数为$F(x,y)$，边缘分布函数分别为$F_X(x)$，$F_Y(y)$，如果对任意的实数x,y都有 $F(x,y)=F_X(x)\cdot F_Y(y)$ [$x\in R, y\in R$，即事件$\{X\leq x\}$与$\{Y\leq y\}$相互独立] 则称X与Y相互独立，否则称X与Y不相互独立.
  - 相互独立充要条件
    - ① 离散型随机变量X和Y相互独立的充要条件：对任意i,j=1,2,…，有
      $P\{X=x_i, Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}$ ⟺ $p_{ij}=p_{i\cdot}\cdot p_{\cdot j}$
    - ② 连续型随机变量X和Y相互独立的充要条件：对任意的x,y，有
      $f(x,y)=f_X(x)\cdot f_Y(y)$

相互独立的性质
- ①、设 $X_1, X_2, \cdots, X_n$ 相互独立，则其中任意 $k(2\le k\le n)$ 个随机变量也相互独立
- ②、设 (X,Y) 为二维离散型随机变量，X与Y独立，则条件分布等于边缘分布．
- ③、设 (X,Y) 为二维连续随机变量，X与Y独立，则条件概率密度等于边缘概
- ④、若 $X_1, X_2, \cdots, X_n$ 相互独立，$g(x_1), g(x_2), \cdots, g(x_n)$ 为一元连续函数，则 $g($

多维随机变量函数的分布
- 概念：设X,Y为随机变量，$g(x,y)$ 为二元函数，则以随机变量X,Y作为变量的函数 $U=g(X,Y)$ 变量，称之为随机变量X,Y的函数．
- 求法[已知(X,Y)的联合分布求，$Z=g(X,Y)$]
  - ①若(X,Y)是二维离散型随机变量，则 $Z=g(X,Y)$ 也是
    然后根据(X,Y)的联合概率分布即可求得Z的
  - ②如果X,Y其中一个是离散型的，另一个是非离散型
    然后用全概率公式即可得到Z的分布．
  - ③．如果(X,Y)是二维连续型随机变量，即 $(X,Y)\sim f(x,y)$，
    $$F_Z(z)=P\{Z\le z\}=P\{g(X,Y)\le z\}=\iint\limits_{g(x,y)\le z} f(x,y)$$
- 相互独立随机变量函数的分布及卷积公式
  - ①和的分布：$(X,Y)\sim f(x,y)$，则 $Z=X+Y$ 的概率密度为
    证：$F_Z(z)=P\{Z\le z\}=P\{X+Y\le z\}=\iint\limits_{x+y\le z} f(x,y)dxdy=\int_{-\infty}^{+\infty}dy$
    $\Rightarrow f_Z(z)=[F_Z(z)]'=\left\{\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-y} f(x,y)\,dx\right]dy\right\}'_z=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-y}\right.$
    $=\int_{-\infty}^{+\infty} f(z-y,y)\,dy$
    同理 $=\int_{-\infty}^{+\infty} f(x,z-x)\,dx$
    $\Rightarrow$ 当X与Y相互独立时，有卷积
  - ②差的分布：设 $(X,Y)\sim f(x,y)$，则 $Z=X-Y$ 的概率密度为
    证：$F_Z(z)=P\{Z\le z\}=P\{X-Y\le z\}=\iint\limits_{x-y\le z} f(x,y)dx=\int_{-\infty}^{+\infty}$
    $\Rightarrow f_Z(z)=[F_Z(z)]'=\left\{\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z+y} f(x,y)\,dx\right]dy\right\}'_z=\int_{-\infty}^{+\infty}$
    $=\int_{-\infty}^{+\infty} f(z+y,y)\,dy$
    同理 $=\int_{-\infty}^{+\infty} f(x,x-z)\,dx$
  - 其他分布计算类似……

设随机变量X与Y相互独立，则：
①若 $X\sim B(n,p)$，$Y\sim B(m,p)$ $\Rightarrow$ $X+Y\sim B(n+m,p)$

常见分布的可加性

② 若 $X \sim P(\lambda_1)$，$Y \sim P(\lambda_2)$ $\Rightarrow$ $X+Y \sim P(\lambda_1+\lambda_2)$

③ 若 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$ $\Rightarrow$ $X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ [$aX+bY \sim N(a\mu_1+b\mu_2, a^2\sigma_1^2+b^2\sigma_2^2)$]

④ 若 $X \sim \chi^2(n)$，$Y \sim \chi^2(m)$ $\Rightarrow$ $X+Y \sim \chi^2(n+m)$.