期末复习: 概率篇

2023 年 2 月 19 日

目录

1 事件及其概率

1.1 事件

事件定义和基本运算

1.2 概率

1. 古典概型(事件的计数, 一般是排列组合, 可能用到事件的运算)
2. 条件概率

$$P(A|B)=\frac{P(AB)}{P(B)} \tag{1}$$

3. 全概率公式 对完备事件群$\{B_i\}_{i=1}^n$,

$$P(A)=\sum_{i=1}^n P(A|B_i)P(B_i), \tag{2}$$

4. 贝叶斯公式

$$P(B_i|A)=\frac{P(B_iA)}{P(A)}=\frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}, \tag{3}$$

5. 独立 n个事件相互独立,

$$P(\prod_{i=1}^n A_i)=\prod_{i=1}^n P(A_i). \tag{4}$$

   两两独立 vesus 相互独立
6. 书上1.6节给了一些进阶求概率的方法, 但是一般比较难.

2 随机变量及其分布

2.1 离散型随机变量及其分布

1. 0-1分布/伯努利分布

$$P(X=1)=1-P(X=0)=p, E(X)=p, \mathrm{Var}(X)=p(1-p).$$

2. 离散均匀分布 $P(X=x_k)=1/n,\ k=1,\cdots,n.$
3. 二项分布$B(n,p)$ n次实验某事件发生的次数, n个独立的0-1分布的随机变量和的分布

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\ E(X)=np,\ \mathrm{Var}(X)=np(1-p). \tag{5}$$

4. 负二项分布$NB(r,p)$ 重复直到某事件发生r次时的实验次数

$$P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r},\ E(X)=r/p,\ \mathrm{Var}(X)=\frac{r(1-p)}{p^2},$$

如果取$r=1$, 则$X\sim Ge(p)$, 因此负二项分布可以看作r个独立的几何分布随机变量和.

5. 泊松分布$P(\lambda)$ 二项分布的极限

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!},\ E(X)=\lambda,\ \mathrm{Var}(X)=\lambda. \tag{6}$$

2.2 连续型

1. 分布函数和概率密度函数
2. 均匀分布$U(a,b)$

$$f(x)=\frac{1}{b-a}I(a<x<b),\ E(X)=\frac{a+b}{2},\ \mathrm{Var}(X)=\frac{(b-a)^2}{12}. \tag{7}$$

3. 指数分布$Exp(\lambda)$ 刻画寿命

$$f(x)=\lambda e^{-\lambda x}I(x>0),\ E(X)=\frac{1}{\lambda},\ \mathrm{Var}(X)=\frac{1}{\lambda^2}. \tag{8}$$

独立指数分布变量和与卡方分布以及伽马分布的关系.

4. 正态分布$N(\mu,\sigma^2)$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),\ E(X)=\mu,\ \mathrm{Var}(X)=\sigma^2. \tag{9}$$

正态分布的线性变换依然为正态分布, 独立标准正态的平方和服从卡方分布.

2.3 随机变量函数的分布

计算随机变量函数在值域上的概率, 可以转化为求随机变量在定义域的概率. 密度变换公式 $Y=g(X),\ X\sim f$, h为g的反函数.

$$f(y)=f(h(y))|h'(y)|I(-\infty<h(y)<\infty). \tag{10}$$

3 多维随机变量及其分布

3.1 多维随机变量及其分布

1. 多维离散型随机变量的联合概率质量函数, 如多项分布$M(N; p_1, p_2, \cdots, p_n)$

$$P(X_1 = k_1, \cdots, X_n = k_n) = \frac{N!}{k_1!k_2!\cdots k_n!} p_1^{k_1} \cdots p_n^{k_n} \text{ for } k_1 + k_2 + \cdots + k_n = N, \tag{11}$$

多项分布的其中几项的分布依然为多项分布, 边缘分布为二项分布.

2. 多维连续性随机变量的联合密度函数, 如二元正态分布$X_1, X_2 \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, $\mathrm{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2$, 边缘分布$X_i \sim N(\mu_i, \sigma_i^2)$.

3. 多维联合密度的密度转换公式可见书上106页

$$p(\mathbf{y}) = f(\mathbf{g}^{-1}(\mathbf{y}))|J|I_D(\mathbf{y}).$$

3.2 边缘分布

对冗余变量求和或者求积分, 如

$$P(X_1 = i) = \sum_{j=1}^{n} P(X_1 = i, X_2 = j)$$
$$f(x) = \int f(x, y)dy$$
$$F(x) = \lim_{y \to \infty} F(x, y)$$

3.3 条件密度与独立随机变量

1. 条件密度

$$f(y|x) = \frac{f(x, y)}{f_X(x)} \tag{12}$$

2. 独立随机变量的联合分布函数或者联合密度函数为边缘密度或分布的乘积.

3.4 随机向量函数的分布

常见如随机变量和,商与次序统计量的分布, 可以先在坐标轴上画出符合条件的区域求分布函数, 再对分布函数求导得到密度; 也可以构造一组简

单的变量, 用密度变换公式求联合密度, 再求边缘分布, 如求$Z = X + Y$, 可以先求(Z, Y)的联合密度, 再对Y积分.

4 数字特征和极限定理

4.1 数字特征

1. 期望 $EX = \int xg(x)dx$ or $EX = \sum_x xP(X = x)$.
   (1) 线性性(不需要独立);
   (2) 对独立随机变量, $E(X_1X_2) = E(X_1)E(X_2)$;
   (3) 随机变量函数的期望 $Ef(X) = \int f(x)g(x)dx$ for $X \sim g$.
2. 条件期望

$$E[Y|X = x] = \sum_y yP(Y = y|X = x),$$
$$E[Y|X = x] = \int yf_{Y|X}(y|x)dx, \tag{13}$$

   平滑公式 $E[E[Y|X]] = E[Y]$.
3. 中位数和众数
4. 方差, 协方差, 相关系数

$$\begin{aligned}\mathrm{Var}(X) &= E[X - EX]^2 = EX^2 - (EX)^2,\\ \mathrm{Cov}(X, Y) &= E[(X - EX)(Y - EY)] = E(XY) - EXEY,\\ \rho &= \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.\end{aligned}$$

一些性质:

(1) $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$, $\mathrm{Var}(X + c) = \mathrm{Var}(X)$.

(2) 常数方差为0;

(3) 独立随机变量和的方差等于各自方差的和;

(4) $\mathrm{Var}(X) = 0$ iff $P(X = c) = 1$, 对任意常数c, $\mathrm{Var}(X) \leq E(X - c)^2$;

(5) 马尔可夫不等式 $P(Y \geq \epsilon) \leq EY/\epsilon$; 切比雪夫不等式 $P(|Y - \mu| \geq \epsilon) \leq \mathrm{Var(Y)}/\epsilon^2$;

(6) 柯西-施瓦茨不等式 $[Cov(X, Y)]^2 \leq \mathrm{Var}(X)\mathrm{Var}(Y)$, 等号成立当且仅当$X, Y$有严格线性关系.

4.2 大数定律和中心极限定理

1. 大数定律

$$\frac{\sum_{i=1}^{n} X_i}{n} \to EX_1.$$

2. 中心极限定理

$$\frac{S_n/n - \mu}{\sqrt{\mathrm{Var}(X_1)/n}} \to N(0,1).$$