第三次习题课的补充

2022 年 11 月 5 日

1 复习–随机变量的数字特征和极限定理

注意此部分并非本章所有知识点, 也不严谨, 不能作为唯一复习资料, 仅作参考!

1.1 数学期望和中位数

1.1.1 数学期望的定义和性质

假设随机变量X的期望存在, 其期望**定义**为

$$E(X)=\begin{cases}\sum_{k\geq 1} x_k p_k, & X\text{离散};\\ \int xf(x)dx, & X\text{连续}.\end{cases} \tag{1}$$

期望的性质:

(1) **线性性** 若$X_1,\cdots,X_n$为一列随机变量, 它们的期望均存在, $a_1,\cdots,a_n$为常数, 则

$$E(a_1X_1+\cdots+a_nX_n)=a_1E(X_1)+\cdots+a_nE(X_n). \tag{2}$$

(2) 若$X_1,\cdots,X_n$为n个相互独立的随机变量, 它们的期望均存在, 则

$$E(X_1\cdots X_n)=\prod_{k=1}^{n}E(X_k). \tag{3}$$

(3) 随机变量函数的期望 设$\mathbf{X}$为一个n维随机变量, 有分布函数$F_{\mathbf{X}}(\mathbf{x})$,

$\mathbf{Y}=g(\mathbf{X})$为m维随机变量且分布函数维$F_{\mathbf{Y}}(\mathbf{y})$, 若$\mathbf{Y}$的各分量存在期望, 则

$$E(\mathbf{Y})=\int_{\mathbb{R}^m}\mathbf{y}dF_{\mathbf{Y}}(\mathbf{y})=\int_{\mathbb{R}^n}g(\mathbf{x})dF_{\mathbf{X}}(\mathbf{x}) \tag{4}$$

$$=\begin{cases}\sum g(\mathbf{x})P(\mathbf{X}=\mathbf{x}), & \mathbf{X}\text{离散型}\\ \int g(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}, & \mathbf{X}\text{连续性.}\end{cases} \tag{5}$$

1.1.2 条件数学期望

定义 给定$X=x$时随机变量Y的条件期望

$$E(Y|X=x)=\begin{cases}\sum y_kP(Y=y_k|X=x), & \text{离散;}\\ \int yf_{Y|X}(y|x)dy, & \text{连续.}\end{cases} \tag{6}$$

条件期望的平滑公式 设X,Y为随机变量, E表示期望.

$$E[X]=E\{E[X|Y]\}. \tag{7}$$

1.1.3 中位数和众数

中位数 满足

$$P(X\geq m)\geq 1/2,\ P(X\leq m)\geq 1/2$$

的常数m为随机变量X的中位数.

众数 使得概率质量函数(离散)或者概率密度函数(连续)最大的值称为众数.

p**分位数** 设$0<p<1$, 称Q_p是随机变量X的p分位数, 如果

$$P(X\leq Q_p)\geq p,\ P(X\geq Q_p)\leq 1-p.$$

1.2 方差和矩

1.2.1 方差和标准差

$$\text{方差: } \mathrm{Var}(X)=E[(X-E[X])^2]=E[X^2]-(E[X])^2; \tag{8}$$

$$\text{标准差}=\sqrt{\mathrm{Var}(X)}. \tag{9}$$

方差的性质:

(1) 常数方差为0;

(2) 设c为常数, 则 $\mathrm{Var}(cX)=c^2\mathrm{Var}(X)$, $\mathrm{Var}(X+c)=\mathrm{Var}(X)$;

(3) **独立**随机变量和的方差等于随机变量方差的和, 即若$X_1,\cdots,X_n$相互独立, 则

$$\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathrm{Var}(X_i);$$

(4) $\mathrm{Var}(X)=0$当且仅当$P(X=c)=1$, 其中$c=E(X)$. 对任何常数c有, $\mathrm{Var}(X)\leq E(X-c)^2$, 其中等号成立当且仅当$c=E(X)$.

马尔可夫不等式 若随机变量$Y\geq 0$, 则任意$\epsilon>0$, 有

$$P(Y\geq\epsilon)\leq\frac{E(Y)}{\epsilon}. \tag{10}$$

由马尔可夫不等式可得切比雪夫不等式.

1.2.2 矩

定义 设X为随机变量, 满足$E(|X|^k)<\infty$, k为正整数, 则$E(X-c)^2$称为X关于c的k阶矩, 其中c为常数. 称$\alpha_k=E(X^k)$为随机变量X的k阶原点矩, 称$\mu_k=E(X-E(X))^k$为X的k阶中心矩.

矩母函数的定义和定理4.2见书上. 证明两个随机变量分布相同, 可利用矩母函数来证明, 也可以利用特征函数来证明.

1.2.3 协方差和相关系数

协方差 $\mathbf{Cov}(X,Y)=E[(X-EX)(Y-EY)]$.

性质:

(1) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$;

(2) $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)$;

(3) 对任意实数a,b,c,d有

$$\begin{aligned}&\mathrm{Cov}(aX+b,cY+d)=ac\mathrm{Cov}(X,Y),\\&\mathrm{Cov}(aX+bY,cX+dY)=ac\mathrm{Var}(X)+(ad+bc)\mathrm{Cov}(X,Y)+bd\mathrm{Var}(Y).\end{aligned}$$

(4) (a)若X,Y相互独立, 则$\mathrm{Cov}(X,Y)=0$.

(b) $[\mathrm{Cov}(X,Y)]^2\leq\mathrm{Var}(X)\mathrm{Var}(Y)$. 等号成立当且仅当$X,Y$之间有严格的线性关系.

相关系数

$$\rho_{X,Y} = \mathrm{Cov}\left(\frac{X-E(X)}{\sqrt{\mathrm{Var}(X)}}, \frac{Y-E(Y)}{\sqrt{\mathrm{Var}(Y)}}\right) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} \tag{11}$$

正态随机变量独立等价于不相关.

1.3 熵

$$H(X) = \begin{cases} -\sum_{k=1}^{\infty} p_k \log_2 p_k, & X\text{离散}; \\ -\int_{-\infty}^{\infty} f_X(x) \ln f_X(x) dx, & X\text{连续}. \end{cases} \tag{12}$$

关注一下书上熵的性质和例题

1.4 大数定律和中心极限定理

依概率收敛 $\lim_{n\to\infty} P(|X_n - X|\geq\epsilon) = 0$
依分布收敛 $\lim_{n\to\infty} P(X_n\leq x) = P(X\leq x)$
依概率收敛和依分布收敛的关系
大数定律 $\sum_{i=1}^{n} X_i/n \xrightarrow{P} E(X_1)$
中心极限定理

$$\frac{\sqrt{n}(\sum_{i=1}^{n} X_i/n - EX_1)}{\sqrt{\mathrm{Var}(X_1)}} \xrightarrow{\mathcal{L}} N(0,1) \tag{13}$$

2 一些补充

1. Jensen不等式

我们首先定义凸函数, 称函数f为凸函数, 如果对任意x_1, x_2和$0 < t < 1$有

$$tf(x_1) + (1-t)f(x_2)\geq f(tx_1 + (1-t)x_2).$$

Jensen不等式 若f为凸函数, 则有

$$E[f(X)]\geq f(E[X]). \tag{14}$$

如$f(x) = x^2$或者$|x|$.

2. 定义给定Y时X的条件方差

$$\mathrm{Var}(X|Y) = E[X^2|Y] - (E[X|Y])^2, \tag{15}$$

注意此条件方差为Y的函数, 也为一个随机变量.

有如下公式,

$$\mathrm{Var}(X) = E[\mathrm{Var}(X|Y)] + \mathrm{Var}(E[X|Y]). \tag{16}$$

证明:

我们从等式右边(Right Hand Side, RHS)推左边(Left Hand Side, LHS),

$$\begin{aligned}
\mathrm{RHS} =& E[\mathrm{Var}(X|Y)] + \mathrm{Var}(E[X|Y]) \\
=& E\{E[X^2|Y]\} - E[(E[X|Y])^2] + E[(E[X|Y])^2] - (E\{E[X|Y]\})^2 \\
=& E[X^2] - (E[X])^2 \\
=& \mathrm{Var}(X) = \mathrm{LHS},
\end{aligned}$$

其中第二个等号是由于条件方差和方差的定义, 第三个等号用到条件期望的平滑公式. ■

3. 记$\mathbf{X} = (X_1, \cdots, X_n)^\top$表示$n$维随机向量, 里面的每个元素$X_i$都是一个随机变量. 随机向量$\mathbf{X}$的协方差为下面的$n \times n$对称矩阵

$$\mathrm{Cov}(\mathbf{X}) = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} \tag{17}$$

其中对$i = 1, \cdots, n$有$v_{ii} = \mathrm{Var}(X_i)$, 对$i \neq j$有$v_{ij} = \mathrm{Cov}(X_i, X_j)$.

例如对二元正态分布$\mathbf{X} = (X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$,

$$\mathrm{Cov}(\mathbf{X}) = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

设$\mathbf{Y} = M_{m\times n}\mathbf{X}$为$m$维随机向量, 则

$$\mathrm{Cov}(\mathbf{Y}) = \mathrm{Cov}(M_{m\times n}\mathbf{X}) = M_{m\times n}\mathrm{Cov}(\mathbf{X})M_{m\times n}^\top. \tag{18}$$