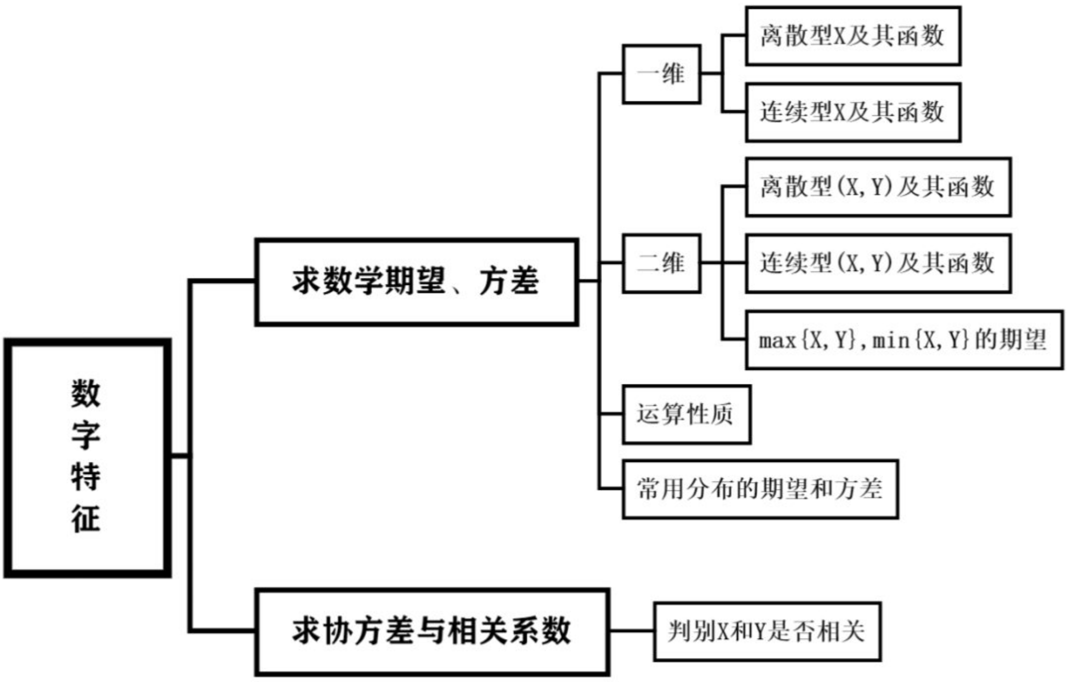
第四章 数字特征

知识图解



知识讲解

一、数学期望

1. 定义

离散型 设离散型随机变量 X 的分布律为 $P\{X=x_k\}=p_k\ (k=1,2,\cdots)$，若级数 $\sum_{k=1}^{\infty}x_k p_k$ 绝对收敛，则称该级数为 X 的**数学期望**，记为 EX，即

$$EX=\sum_{k=1}^{\infty}x_k p_k$$

否则，称 X 的数学期望不存在.

注：数学期望 EX 的唯一性要求与 X 的取值顺序无关，即 $\sum_{k=1}^{\infty}x_k p_k$ 绝对收敛. 故若 x_k 为有限个，则数学期望 EX 一定存在.

连续型 设连续型随机变量 X 的密度函数为 ，若 $\int_{-\infty}^{+\infty}|x|f(x)dx<+\infty$，则称

$$EX=\int_{-\infty}^{+\infty}xf(x)dx$$

为 X 的数学期望，否则称 X 的数学期望不存在.

2. 随机变量函数的期望

1）一维随机变量的函数 $Y=g(X)$ 的期望

离散型 X 设离散型随机变量 X 的分布律为 $P\{X=x_k\}=p_k\ (i=1,2,\cdots)$，则它的函数 $Y=g(X)$ 的数学期望为（若 $\sum\limits_{k=1}^{\infty}g(x_k)p_k$ 绝对收敛）

$$EY=E[g(X)]=\sum_{k=1}^{\infty}g(x_k)p_k$$

连续型 X 设连续型随机变量 X 的密度函数为 $f(x)$，则它的函数 $Y=g(X)$ 的数学期望为（若 $\int\limits_{-\infty}^{+\infty}g(x)f(x)dx$ 绝对收敛）

$$EY=E[g(X)]=\int_{-\infty}^{+\infty}g(x)f(x)dx$$

【例4.1】 设随机变量 X 的分布律为 $X\sim\begin{pmatrix}-2 & 0 & 2\\ \frac{4}{10} & \frac{3}{10} & \frac{3}{10}\end{pmatrix}$，求 $EX,E(3X^2+5)$.

【例4.2】 设随机变量 X 的概率密度为 $f(x)=\begin{cases} e^{-x}, & x>0 \\ 0, & x\le 0 \end{cases}$，求：

（1）EX；（2）$Y=e^{-2X}$ 的数学期望 EY；（3）$Z=\max\{X,1\}$ 的数学期望 EZ .

【例4.3】 （练习）设随机变量 X 的概率密度为 $f(x)=\begin{cases} \frac{3}{8}x^2, & 0<x<2, \\ 0, & 其他. \end{cases}$ 则

$E\left(\frac{1}{X^2}\right)=$__________.

【例4.4】 设随机变量 X 的概率密度为 $f(x)=\frac{1}{\pi\left(1+x^2\right)},-\infty<x<+\infty$，求 EX .

2. 数学期望的性质：

1） $E(C)=C$（其中 C 为常数）

2） $E(X+C)=E(X)+C$

3） $E(CX)=CE(X)$

4） $E(X\pm Y)=E(X)\pm E(Y)$

5）设 X,Y 相互独立，则 $E(XY)=EXEY$

推论 1 $E(c_1X_1+c_2X_2+\cdots c_nX_n+b)=$________________________________.

其中 $c_1,c_{2,}\cdots,c_n,b$ 均是常数.

推论 2 若 $X_1,X_2,\cdots,X_n$ 相互独立，$E(X_1\cdot X_2\cdot\cdots\cdot X_n)=$______________________.

2）二维随机变量函数 $Z=g(X,Y)$ 的期望

离散型 (X,Y) 设二维离散型随机变量 (X,Y) 的分布律为

$P\{X=x_i,Y=y_j\}=p_{ij}$，$i,j=1,2,\cdots$，则函数 $Z=g(X,Y)$ 的数学期望为（若 $\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}g(x_i,y_i)p_{ij}$ 绝对收敛）

$$EZ=E[g(X,Y)]=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}g(x_i,y_i)p_{ij}$$

连续型 (X,Y) 设二维连续型随机变量 (X,Y) 的概率密度为 $f(x,y)$,则

$EZ=E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy$（要求右端积分绝对收敛）.

【例4.5】　设随机变量(X,Y)的分布律为

$Y \backslash X$	1	2	3
-1	0.2	0.1	0.0
0	0.1	0.0	0.3
1	0.1	0.1	0.1

设$Z=(Y-X)^2$，求EZ

【例4.6】　设随机变量(X,Y)的概率密度

$$f(x,y)=\begin{cases}\dfrac{3}{2x^3y^2}, & \dfrac{1}{x}<y<x,x>1\\ 0, & \text{其他}\end{cases}$$

求数学期望$E\left(\dfrac{1}{XY}\right)$.

【例4.7】　甲、乙两人约定在8~12点某地会面，设两人8点X与Y（小时）到达会面地，两人到达时间相互独立且均服从$[0,4]$上的均匀分布，求先到者的平均等待时间.

二、方差

1. 定义： 设 X 为一随机变量，若 $E(X-EX)^2$ 存在，则称 $E(X-EX)^2$ 为 X 的方差，记为 DX，即

$$DX=E(X-EX)^2$$

称 $\sqrt{DX}$ 为 X 的**标准差**. 易得计算公式 $DX=EX^2-(EX)^2$. 显然 $DX\geq 0$.

离散型 设离散型随机变量 X 的分布律为 $P\{X=x_k\}=p_k,(i=1,2,\cdots)$，则

$$DX=\sum_{k=1}^{\infty}(x_k-EX)^2 p_k \text{（要求右端级数绝对收敛）}$$

连续型 设连续型随机变量 X 的密度函数为 $f(x)$，则

$$DX=\int_{-\infty}^{\infty}(x-EX)^2 f(x)dx. \text{（要求右端积分绝对收敛）}$$

2. 方差的性质：

1） $D(C)=0$（其中 C 为常数）;

2） $D(X+C)=DX$；

3） $D(CX)=C^2DX$；

4）设 X,Y 相互独立，则 $D(X\pm Y)=DX+DY$；

5） $DX=0$ 的充要条件是 $P\{X=EX\}=1$.（证明由切比雪夫不等式）

推 论 若 $X_1,X_2,...,X_n$ 相互独立，则 $D(c_1X_1+c_2X_2+\cdots+c_nX_n+b)=$ ________________.

【例4.8】 设 X 服从正态分布 $N(1,4)$，Y 服从参数为 1 的泊松分布，且 X 与 Y 相互独立，则 $E(XY+1-Y)=$ _______, $D(2Y-X+1)=$ _______.

【例4.9】 设随机变量 X 服从 $\left[-3,2\right]$ 上的均匀分布，随机变量 $Y=\begin{cases}1, & X\geq 0\\ -1, & X<0\end{cases}$，则 $DY=$（ ）

（A）$\frac{1}{5}$ （B）$\frac{1}{25}$ （C）$\frac{24}{25}$ （D）1

【例4.10】 设随机变量 X 的概率密度为 $f\left(x\right)=\begin{cases}1+x, & -1<x<0,\\ 1-x, & 0<x<1,\\ 0, & 其他.\end{cases}$ 则方差 $DX=$ ______.

3. 常用分布的期望和方差：

分布类型	分布律或概率密度	期望	方差
0-1 分布	$p_k=P\{X=k\}=p^kq^{1-k} \quad (q=1-p),(k=0,1)$	p	pq
二项分布	$p_k=P\{X=k\}=C_n^kp^kq^{n-k}$ $(q=1-p),(k=0,1,...,n)$	np	npq
泊松分布	$p_k=P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda} \quad (i=0,1,2...)$	λ	λ
均匀分布	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & 其他\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
正态分布	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (-\infty<x<+\infty,\sigma>0)$	μ	σ^2
指数分布	$f(x)=\begin{cases}\lambda e^{-\lambda x},x>0\\ 0,x\le 0\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
χ^2 分布	$X_1,X_2,...X_n$相互独立，且都服从标准正态分布$N(0,1)$ $\chi^2=X_1^2+X_2^2+...+X_n^2$	n	$2n$
几何分布	$p_k=P\{X=k\}=(1-p)^{k-1}p$， $(k=1,2,...),0<p<1$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$

【例4.11】 设一次试验成功的概率为 p ，进行 100 次独立重复试验，当 $p=$__________时，成功次数的标准差的值最大，其最大值为__________.

【例4.12】 已知随机变量 X 服从二项分布,且 $EX=2.4,DX=1.44$ ，则二项分布的参数 n,p 的值分别为（ ）.

A. $n=4,p=0.6$ B. $n=6,p=0.4$ C. $n=8,p=0.3$ D. $n=24,p=0.1$

【例4.13】 设X服从参数为λ的泊松分布，且$EX^2=6$，则$\lambda=$________.

【例4.14】 设随机变量 X 服从参数为λ的指数分布，则$P\{X>\sqrt{DX}\}=$________.

【例4.15】 设随机变量X服从区间$\left[a,b\right]$上的均匀分布，且$EX=3,DX=3,$则区间$\left[a,b\right]$为________.

【例4.16】 已知随机变量$X\sim N\left(-3,1\right),Y\sim N\left(2,1\right)$，且$X,Y$相互独立，设随机变量$Z=X-2Y+7$，求$Z$的分布.

【例4.17】 （练习）设随机变量$X\sim N\left(1,4\right),Y\sim U\left(0,4\right)$且$X,Y$相互独立，则$D\left(2X-3Y\right)=$（ ）

（A）8 （B）18 （C）24 （D）28

【例4.18】 设随机变量 X 与 Y 相互独立，且都服从参数为 1 的指数分布，记 $U=\max\left\{X,Y\right\},V=\min\left\{X,Y\right\}$，求（I） EU；（II） EV；（III） EUV

结论： $U+V=X+Y,UV=XY,U-V=\left|X-Y\right|$

其中， $U=\max\left\{X,Y\right\}=\dfrac{X+Y+\left|X-Y\right|}{2}$, $V=\min\left\{X,Y\right\}=\dfrac{X+Y-\left|X-Y\right|}{2}$.

三、切比雪夫不等式

定义 设随机变量具有数学期望 $EX=\mu$，方差 $DX=\sigma^2$，则对于任意正数 ε，有下列切比雪夫不等式成立

$$P\left\{\left|X-EX\right|\geq\varepsilon\right\}\leq\frac{DX}{\varepsilon^2}\ \text{或}\ P\left\{\left|X-EX\right|<\varepsilon\right\}\geq1-\frac{DX}{\varepsilon^2}$$

证明：只证连续型随机变量 X ，设其概率密度 $f\left(x\right)$. 有

$$P\left\{\left|X-EX\right|\geq\varepsilon\right\}=\int\limits_{\left|x-EX\right|\geq\varepsilon}f\left(x\right)dx\leq\int\limits_{\left|x-EX\right|\geq\varepsilon}\frac{\left(x-EX\right)^2}{\varepsilon^2}f\left(x\right)dx$$

$$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}\left(x-EX\right)^2f\left(x\right)dx=\frac{DX}{\varepsilon^2}$$

【例4.19】 设随机变量 X 的方差为 2，则根据切比雪夫不等式有估计 $P\{|X-EX|\geq2\}\leq$__________.

【例4.20】 设随机变量 X, Y 的数学期望都是 2，方差分别为 1 和 4，而相关系数为 0.5，则根据切比雪夫不等式 $P\left\{\left|X-Y\right|\geq6\right\}\leq$___________.

四、协方差

1. 协方差：

定义 对于二维随机变量(X,Y)，若$E(X-EX)(Y-EY)$存在，则称它为X与Y的**协方差**，记为$Cov(X,Y)$，即

$$Cov(X,Y)=E(X-EX)(Y-EY).$$

易得计算公式$Cov(X,Y)=EXY-EXEY$.

性质

1）$Cov(X,Y)=Cov(Y,X)$；

2）$Cov(X,C)=0$；

3）$Cov(aX,bY)=abCov(X,Y)$；

4）$Cov(X,X)=DX$；

5）$Cov(X\pm Y,Z)=Cov(X,Z)\pm Cov(Y,Z)$；

6）设X,Y相互独立，则$Cov(X,Y)=0$；

7）$D(X\pm Y)=DX+DY\pm 2Cov(X,Y)$.

证明：

$$D\left(X\pm Y\right)=E\left[\left(X\pm Y\right)-E\left(X\pm Y\right)\right]^2=E\left[\left(X-EX\right)\pm\left(Y-EY\right)\right]^2$$
$$=E\left(X-EX\right)^2+E\left(Y-EY\right)^2\pm 2E\left(X-EX\right)\left(Y-EY\right)=DX+DY\pm 2Cov\left(X,Y\right)$$

【例4.21】 设随机变量 X 和 Y 的联合概率分布为

X \ Y	-1	0	1
0	0.07	0.18	0.15
1	0.08	0.32	0.20

则 X^2 和 Y^2 的协方差 $Cov\left(X^2,Y^2\right)=$ ________.

【例4.22】 设随机变量 $\left(X,Y\right)$ 的密度函数 $f(x,y)=\begin{cases}3, & (x,y)\in G\\ 0, & 其他\end{cases}$，其中区域 G 由曲线 $y=x^2$ 与 $x=y^2$ 围成，求 $Cov\left(X,Y\right)$.

【例4.23】 设 $X_1,X_2,\cdots,X_n(n>2)$ 独立同分布，其方差 $\sigma^2>0$，令 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，则下列选项正确的是（　　）.

（A）$Cov\left(X_1,\bar{X}\right)=\frac{\sigma^2}{n}$　　（B）$Cov\left(X_n,\bar{X}\right)=\sigma^2$

（C）$D\left(X_1+\bar{X}\right)=\frac{n+1}{n}\sigma^2$　　（D）$D\left(X_1-\bar{X}\right)=\frac{n+1}{n}\sigma^2$

2. 相关系数：

定义 对于二维随机变量(X,Y)，若$DX\neq 0,DY\neq 0,$则称

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{DX}\cdot\sqrt{DY}}$$

为X和Y的**相关系数**. 若$\rho_{XY}=0$，称X和Y**不相关**.

性质

1）$|\rho_{XY}|\leq 1,\ \rho_{XY}=\rho_{YX},\ \rho_{XX}=1$；

2）$|\rho_{XY}|=1$的充要条件是X与Y以概率 1 线性相关，即存在常数$a\neq 0$和$b,$有

$$P\{Y=aX+b\}=1$$

3）$|\rho_{XY}|$越接近于 1， 表明X和Y的线性相关程度越大.

注： 协方差与相关系数的区别（举例）——

设二维随机变量(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\frac{8}{3}, & 0<x-y<0.5, 0<x,y<1,\\ 0, & 其他.\end{cases}$$

计算得到$DX=DY=\frac{37}{648},Cov(X,Y)=0.047,\rho_{XY}=0.824$. 说明$X$与$Y$有较强的正相关关系，但协方差很小. 相关系数能更直接反映X,Y相关关系的强弱.

【例4.24】 将长度为$1m$的木棒随机截成两段，则两段长度的相关系数为（ ）

（A）1 （B）$\frac{1}{2}$ （C）$-\frac{1}{2}$ （D）-1.

【例4.25】 某箱装有 100 件产品，其中一、二、三等品分别为 80，10，10 件. 现从中随机抽取一件，记

$$X_i=\begin{cases}1, & \text{抽到}i\text{等品},\\ 0, & \text{其他},\end{cases}\quad (i=1,2,3)$$

求 X_1 与 X_2 的相关系数.（接【例 3.5】）

3. 常用结论

1）任意随机变量 X 与 Y 相互独立是 X 与 Y 不相关的充分不必要条件.

2）对随机变量 X,Y，有下列等价命题：

$$\rho_{XY}=0 \Leftrightarrow Cov(X,Y)=0 \Leftrightarrow EXY=EXEY \Leftrightarrow D(X\pm Y)=DX+DY$$

3）(X,Y) 服从**二维正态分布**，有下列等价命题：

$$X\text{与}Y\text{独立} \Leftrightarrow \rho_{XY}=0 \Leftrightarrow Cov(X,Y)=0 \Leftrightarrow EXY=EXEY \Leftrightarrow D(X\pm Y)=DX+DY$$

【例4.26】 设随机变量X与Y的期望和方差存在，且$D(X-Y)=DX+DY$，则下列说法错误的是（ ）.

（A） $D(X+Y)=DX+DY$ （B）$EXY=EXEY$

（C） X与Y不相关 （D）X与Y相互独立

【例4.27】 （练习）设随机变量X,Y不相关，其概率分布分别为

X	0	3
p	0.6	0.4

Y	-1	0
p	0.7	0.3

则随机事件$\{X=0\}$与$\{Y=0\}$（ ）.

（A）互不相容； （B）相互独立； （C）相互对立； （D）不独立.

【例4.28】 设二维随机变量(X,Y)服从$N(\mu,\mu;\sigma^2,\sigma^2;0)$，则$E(XY^2)=$ ________.