概率论与数理统计

庄玮玮
weizh@ustc.edu.cn

安徽 合肥

2022 年 10 月

第三章 随机变量的数字特征

§3.1 数学期望

§3.1 数学期望

随机变量的分布函数或概率密度描述了随机变量的统计性质，从中可以了解随机变量落入某个区间的概率，但是还不能给人留下更直接的总体印象. 例如用X表示某计算机软件的使用寿命，当知道X服从指数分布$\mathcal{E}(\lambda)$后，我们还不知道该软件的平均使用寿命是多少.

这里的平均使用寿命应当是一个常数. 我们需要为随机变量X定义一个平均值，这就是数学期望,它反映随机变量的平均取值.

§3.1 数学期望

§3.1 数学期望

随机变量的分布函数或概率密度描述了随机变量的统计性质，从中可以了解随机变量落入某个区间的概率，但是还不能给人留下更直接的总体印象. 例如用X表示某计算机软件的使用寿命，当知道X服从指数分布$\mathcal{E}(\lambda)$后，我们还不知道该软件的平均使用寿命是多少.

这里的平均使用寿命应当是一个常数. 我们需要为随机变量X定义一个平均值，这就是数学期望，它反映随机变量的平均取值.

§3.1 数学期望

例3.1.1 甲每次投资成功的概率为70%，失败的概率为30%．假设每次投资成功将获利3万元，投资失败将损失1万元，下次投资时甲期望赢利多少？

解 在以后的n次独立重复的投资中，甲大约有$0.7n$次获利3万元，大约有$0.3n$次损失1万元，每次投资平均获利是

$$\frac{3\times 0.7n-1\times 0.3n}{n}=3\times 0.7-1\times 0.3=1.8.$$

于是，甲下次投资期望获利1.8万元．

§3.1 数学期望

例3.1.1　甲每次投资成功的概率为70%，失败的概率为30%. 假设每次投资成功将获利3万元，投资失败将损失1万元，下次投资时甲期望赢利多少？

解　在以后的n次独立重复的投资中，甲大约有$0.7n$次获利3万元，大约有$0.3n$次损失1万元，每次投资平均获利是

$$\frac{3 \times 0.7n - 1 \times 0.3n}{n} = 3 \times 0.7 - 1 \times 0.3 = 1.8.$$

于是，甲下次投资期望获利1.8万元.

§3.1 数学期望

例3.1.1 甲每次投资成功的概率为70%，失败的概率为30%. 假设每次投资成功将获利3万元，投资失败将损失1万元，下次投资时甲期望赢利多少?

解 在以后的n次独立重复的投资中，甲大约有$0.7n$次获利3万元，大约有$0.3n$次损失1万元，每次投资平均获利是

$$\frac{3\times 0.7n-1\times 0.3n}{n}=3\times 0.7-1\times 0.3=1.8.$$

于是，甲下次投资期望获利1.8万元.

§3.1 数学期望

在上面的例子中，甲的期望值是多次投资的平均收益．如果用随机变量

$$X = \begin{cases} 3, & \text{当投资成功,} \\ -1, & \text{当投资失败} \end{cases}$$

描述甲的投资情况，则称1.8是X的数学期望．用$\mathrm{E}(X)$表示X的数学期望时，有

$$\mathrm{E}(X) = 3P(X=3) + (-1)P(X=-1) = 1.8.$$

例1.1中，数学期望指多次独立重复投资时，每次投资的平均收益．

§3.1 数学期望

在上面的例子中，甲的期望值是多次投资的平均收益. 如果用随机变量

$$X = \begin{cases} 3, & \text{当投资成功}, \\ -1, & \text{当投资失败} \end{cases}$$

描述甲的投资情况，则称1.8是X的数学期望. 用$\mathrm{E}(X)$表示X的数学期望时，有

$$\mathrm{E}(X) = 3P(X = 3) + (-1)P(X = -1) = 1.8.$$

例1.1中，数学期望指多次独立重复投资时，每次投资的平均收益.

§3.1 数学期望

例3.1.2 一个班有$m=180$个学生, 期中考试后有m_j个人的成绩是j分$(0\leqslant j\leqslant 100)$. 成绩是$j$分的学生所占的比例是$p_j=m_j/m$. 用向量

$$(p_0,p_1,p_2,\cdots,p_{100}) \tag{3.1.1}$$

表示这个班期中成绩的分布, 称为总体分布. 用x_i表示第i个学生的成绩, 则期中考试的全班平均分是

$$\mu\equiv\frac{1}{m}\sum_{i=1}^{m}x_i=\frac{1}{m}\sum_{j=0}^{100}j\cdot m_j=\sum_{j=0}^{100}jp_j. \tag{3.1.2}$$

在统计学中, 称全班的分数构成的集合$\{x_j\,|\,1\leqslant j\leqslant m\}$为总体, 称总体的平均$\mu$为总体平均, 而总体的分布(1.1)恰好是总体分布.

§3.1 数学期望

例3.1.2 一个班有$m=180$个学生, 期中考试后有m_j个人的成绩是j分($0\leqslant j\leqslant 100$). 成绩是$j$分的学生所占的比例是$p_j=m_j/m$. 用向量

$$(p_0,p_1,p_2,\cdots,p_{100}) \qquad (3.1.1)$$

表示这个班期中成绩的分布, 称为总体分布. 用x_i表示第i个学生的成绩, 则期中考试的全班平均分是

$$\mu\equiv\frac{1}{m}\sum_{i=1}^{m}x_i=\frac{1}{m}\sum_{j=0}^{100}j\cdot m_j=\sum_{j=0}^{100}jp_j. \qquad (3.1.2)$$

在统计学中, 称全班的分数构成的集合$\{x_j\,|\,1\leqslant j\leqslant m\}$为总体, 称总体的平均$\mu$为总体平均, 而总体的分布(1.1)恰好是总体分布.

§3.1 数学期望

例3.1.2 一个班有$m=180$个学生，期中考试后有m_j个人的成绩是j分$(0\leqslant j\leqslant 100)$. 成绩是$j$分的学生所占的比例是$p_j=m_j/m$. 用向量

$$\left(p_0, p_1, p_2, \cdots, p_{100}\right) \qquad (3.1.1)$$

表示这个班期中成绩的分布，称为总体分布. 用x_i表示第i个学生的成绩，则期中考试的全班平均分是

$$\mu \equiv \frac{1}{m}\sum_{i=1}^{m} x_i = \frac{1}{m}\sum_{j=0}^{100} j\cdot m_j = \sum_{j=0}^{100} jp_j. \qquad (3.1.2)$$

在统计学中，称全班的分数构成的集合$\{x_j \mid 1\leqslant j\leqslant m\}$为总体，称总体的平均$\mu$为总体平均，而总体的分布(1.1)恰好是总体分布.

§3.1　数学期望

现在从班中任选一人，用X表示他的期中成绩，则X有概率分布

$$p_j = P(X = j) = \frac{m_j}{m},\ 0 \leqslant j \leqslant 100. \qquad (3.1.3)$$

因为X是从总体$\{x_j \mid 1 \leqslant j \leqslant m\}$中随机抽样得到的，所以把$X$的数学期望定义成总体平均$\mu$. 用$\mathrm{E}\,(X)$表示$X$的数学期望时，有

$$\mathrm{E}\,(X) = \mu = \sum_{j=0}^{100} j p_j.$$

§3.1 数学期望

现在从班中任选一人，用X表示他的期中成绩，则X有概率分布

$$p_j = P(X=j) = \frac{m_j}{m},\ 0 \leqslant j \leqslant 100. \tag{3.1.3}$$

因为X是从总体$\{x_j \mid 1 \leqslant j \leqslant m\}$中随机抽样得到的，所以把$X$的数学期望定义成总体平均$\mu$. 用$\mathrm{E}(X)$表示$X$的数学期望时，有

$$\mathrm{E}(X) = \mu = \sum_{j=0}^{100} j p_j.$$

§3.1　数学期望

在例3.1.2中，X 的数学期望是总体平均.

设想在班里有放回地独立重复随机选取N次. 当N充分大，因为得j分的学生被选到的概率是p_j，所以被选到的次数大约是Np_j. 这N次随机选择得到的平均分大约是

$$\frac{1}{N}\sum_{j=0}^{100} j \cdot Np_j = \sum_{j=0}^{100} jp_j = \mathrm{E}(X).$$

为此，我们也称$\mathrm{E}(X)$是在班里任选一人时，期望得到的分数.

§3.1 数学期望

在例3.1.2中，X 的数学期望是总体平均.

设想在班里有放回地独立重复随机选取N次. 当N充分大，因为得j分的学生被选到的概率是p_j，所以被选到的次数大约是Np_j. 这N次随机选择得到的平均分大约是

$$\frac{1}{N}\sum_{j=0}^{100} j\cdot Np_j = \sum_{j=0}^{100} jp_j = \mathrm{E}\,(X).$$

为此，我们也称$\mathrm{E}\,(X)$是在班里任选一人时，期望得到的分数.

§3.1　数学期望

A. 数学期望的定义

下面引入随机变量的数学期望的定义.

定义3.1.1　设X有概率分布

$$p_j = P(X = x_j),\ j = 0, 1, \cdots,$$

如果$\sum_{j=0}^{\infty} |x_j| p_j < \infty$, 则称$X$的数学期望存在, 并且称

$$\mathrm{E}(X) = \sum_{j=0}^{\infty} x_j p_j \qquad (3.1.4)$$

为X或分布$\{p_j\}$的数学期望(expected value).

§3.1 数学期望

A. 数学期望的定义

下面引入随机变量的数学期望的定义.

定义3.1.1 设X有概率分布

$$p_j = P(X = x_j),\ j = 0, 1, \cdots,$$

如果$\sum_{j=0}^{\infty} |x_j| p_j < \infty$, 则称$X$的数学期望存在, 并且称

$$\mathrm{E}\,(X) = \sum_{j=0}^{\infty} x_j p_j \qquad (3.1.4)$$

为X或分布$\{p_j\}$的数学期望(expected value).

§3.1 数学期望

在定义3.1.1中，要求$\sum_{j=0}^{\infty}|x_j|p_j<\infty$ 的原因是要使(3.1.4)中的级数有确切的意义. 当所有的x_j非负时，如果(3.1.4)中的级数是无穷，则由(1.4)定义的$\mathrm{E}(X)$也有明确的意义，它表明X的平均取值是无穷. 这时称X的数学期望是无穷.

在定义3.1.1中，将p_j视为$\{p_i\}$在横坐标x_j处的质量，由

$$\sum_{j=1}^{\infty}(x_j-\mu)p_j=\sum_{j=1}^{\infty}x_jp_j-\mu=0,$$

知道$\{p_i\}$的重心是μ. 所以X的数学期望$\mathrm{E}(X)$还是其概率分布$\{p_i\}$的重心.

§3.1 数学期望

在定义3.1.1中，要求$\sum_{j=0}^{\infty}|x_j|p_j<\infty$ 的原因是要使(3.1.4)中的级数有确切的意义．当所有的x_j非负时，如果(3.1.4)中的级数是无穷，则由(1.4)定义的$\mathrm{E}(X)$也有明确的意义，它表明X的平均取值是无穷．这时称X的数学期望是无穷．

在定义3.1.1中，将p_j视为$\{p_i\}$在横坐标x_j处的质量，由

$$\sum_{j=1}^{\infty}(x_j-\mu)p_j=\sum_{j=1}^{\infty}x_jp_j-\mu=0,$$

知道$\{p_i\}$的重心是μ．所以X的数学期望$\mathrm{E}(X)$还是其概率分布$\{p_i\}$的重心．

§3.1 数学期望

对于有概率密度$f(x)$的连续型随机变量X, 我们也用$f(x)$和横轴所夹面积的几何重心定义X的数学期望. 设μ是所述的重心, 如果

$$\int_{-\infty}^{\infty} |x| f(x)\,\mathrm{d}x < \infty, \tag{3.1.5}$$

则有

$$\int_{-\infty}^{\infty} (x-\mu) f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x - \mu = 0.$$

于是$\mu = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$ 是所述的重心.

§3.1 数学期望

对于有概率密度$f(x)$的连续型随机变量X，我们也用$f(x)$和横轴所夹面积的几何重心定义X的数学期望. 设μ是所述的重心，如果

$$\int_{-\infty}^{\infty} |x| f(x)\,\mathrm{d}x < \infty, \qquad (3.1.5)$$

则有

$$\int_{-\infty}^{\infty} (x-\mu) f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x - \mu = 0.$$

于是$\mu = \displaystyle\int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$ 是所述的重心.

§3.1 数学期望

定义3.1.2 设X是有概率密度$f(x)$的随机变量，如果(1.5)成立，则称X的数学期望存在，并且称

$$\mathrm{E}(X)=\int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x \qquad (3.1.6)$$

为X或$f(x)$的数学期望.

和离散时的情况一样，在定义1.2中要求条件(1.5)的原因是要使(1.6)中的积分有确切的意义. 当X非负时，如果(1.6)等于无穷，则由(1.6)定义的$\mathrm{E}(X)$ 也有明确的意义，它表明X的平均取值是无穷. 这时称X的数学期望是无穷.

由于随机变量的数学期望由随机变量的概率分布唯一决定，所以也可以对概率分布定义数学期望. 概率分布的数学期望就是以它为概率分布的随机变量的数学期望. 有相同分布的随机变量必有相同的数学期望.

§3.1 数学期望

定义3.1.2 设X是有概率密度$f(x)$的随机变量，如果(1.5)成立，则称X的数学期望存在，并且称

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x \tag{3.1.6}$$

为X或$f(x)$的数学期望.

和离散时的情况一样，在定义1.2中要求条件(1.5)的原因是要使(1.6)中的积分有确切的意义．当X非负时，如果(1.6)等于无穷，则由(1.6)定义的$\mathrm{E}(X)$也有明确的意义，它表明X的平均取值是无穷．这时称X的数学期望是无穷．

由于随机变量的数学期望由随机变量的概率分布唯一决定，所以也可以对概率分布定义数学期望．概率分布的数学期望就是以它为概率分布的随机变量的数学期望．有相同分布的随机变量必有相同的数学期望.

§3.1 数学期望

定义3.1.2 设X是有概率密度$f(x)$的随机变量，如果(1.5)成立，则称X的数学期望存在，并且称

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x) \, \mathrm{d}x \qquad (3.1.6)$$

为X或$f(x)$的数学期望.

和离散时的情况一样，在定义1.2中要求条件(1.5)的原因是要使(1.6)中的积分有确切的意义. 当X非负时，如果(1.6)等于无穷，则由(1.6)定义的$\mathrm{E}(X)$也有明确的意义，它表明X的平均取值是无穷. 这时称X的数学期望是无穷.

由于随机变量的数学期望由随机变量的概率分布唯一决定，所以也可以对概率分布定义数学期望. 概率分布的数学期望就是以它为概率分布的随机变量的数学期望. 有相同分布的随机变量必有相同的数学期望.

§3.1 数学期望

例3.1.3 17世纪曾有人向帕斯卡请教如下的问题：两个赌博水平相当的人各出50法郎作赌注，并约定五局三胜者获得这100法郎．前三局中甲赢了2局，乙赢了1局，这时因故要中止赌博，问应当怎样合理分配这100法郎．

解 下面是帕斯卡的回答．设想赌博可以继续下去，再赌两局必出结果，这两局的结果只能是以下四个事件之一：

$$\text{甲乙}, \quad \text{乙甲}, \quad \text{甲甲}, \quad \text{乙乙},$$

其中的"甲乙"表示第一局甲胜，第二局乙胜；"乙甲"表示第一局乙胜，第二局甲胜；……．这4个事件发生的可能性相同．因为甲在前3局中赢了2局，所以只有"乙乙"发生时，甲获得0法郎，否则甲获得100法郎．用X表示甲应当分到的赌资，按照以上分析有

$$P(X=0)=P(\text{乙乙})=\frac{1}{4}, \quad P(X=100)=1-\frac{1}{4}=\frac{3}{4}.$$

§3.1 数学期望

例3.1.3 17世纪曾有人向帕斯卡请教如下的问题：两个赌博水平相当的人各出50法郎作赌注，并约定五局三胜者获得这100法郎．前三局中甲赢了2局，乙赢了1局，这时因故要中止赌博，问应当怎样合理分配这100法郎．

解 下面是帕斯卡的回答．设想赌博可以继续下去，再赌两局必出结果，这两局的结果只能是以下四个事件之一：

甲乙， 乙甲， 甲甲， 乙乙，

其中的“甲乙”表示第一局甲胜，第二局乙胜；“乙甲”表示第一局乙胜，第二局甲胜；……．这4个事件发生的可能性相同．因为甲在前3局中赢了2局，所以只有“乙乙”发生时，甲获得0法郎，否则甲获得100法郎．用X表示甲应当分到的赌资，按照以上分析有

$$P(X=0)=P(\text{乙乙})=\frac{1}{4},\quad P(X=100)=1-\frac{1}{4}=\frac{3}{4}.$$

§3.1 数学期望

于是甲期望分得的赌资是

$$\mathrm{E}(X) = 0 + 100 \times \frac{3}{4} = 75(\text{法郎}).$$

这个例子也是“数学期望”的来源之一.

在例3.1.3中，认为甲应当分得赌资的2/3是不合理的. 因为按照这个逻辑，甲赢第一局就因故停止赌博时，甲将获得全部的赌资.

§3.1 数学期望

于是甲期望分得的赌资是

$$\mathrm{E}(X) = 0 + 100 \times \frac{3}{4} = 75(\text{法郎}).$$

这个例子也是“数学期望”的来源之一.

在例3.1.3中，认为甲应当分得赌资的2/3是不合理的. 因为按照这个逻辑，甲赢第一局就因故停止赌博时，甲将获得全部的赌资.

§3.1 数学期望

于是甲期望分得的赌资是

$$\mathrm{E}(X) = 0 + 100 \times \frac{3}{4} = 75(\text{法郎}).$$

这个例子也是“数学期望”的来源之一.

在例3.1.3中，认为甲应当分得赌资的2/3是不合理的. 因为按照这个逻辑，甲赢第一局就因故停止赌博时，甲将获得全部的赌资.

§3.1 数学期望

例3.1.4 在澳门赌场，有很多人在赌廿一点时顺便押对子. 其规则如下：庄家从6副(每副52张)扑克中随机发给你两张. 如果你下注a元，当得到的两张牌是一对时，庄家赔你十倍，否则输掉你的赌注. 如果你下注100元，你在每局中期望赢多少元?

解 用X表示你在一局中的获利，$a = 100$. 则

$$P(X = 10a) = \frac{13\mathrm{C}_{4\times 6}^2}{\mathrm{C}_{52\times 6}^2} = 0.074, \quad P(X = -a) = 1 - 0.074,$$

于是，你期望赢

$$\mathrm{E}(X) = 10a \cdot 0.074 - a \cdot (1 - 0.074) = -0.186a = -18.6(\text{元}).$$

当只使用一副扑克，可以计算出你每局期望赢-35.29元.

§3.1 数学期望

例3.1.4 在澳门赌场，有很多人在赌廿一点时顺便押对子．其规则如下：庄家从6副(每副52张)扑克中随机发给你两张．如果你下注a元，当得到的两张牌是一对时，庄家赔你十倍，否则输掉你的赌注．如果你下注100元，你在每局中期望赢多少元？

解 用X表示你在一局中的获利，$a=100$．则

$$P(X=10a)=\frac{13\mathrm{C}_{4\times 6}^{2}}{\mathrm{C}_{52\times 6}^{2}}=0.074,\quad P(X=-a)=1-0.074,$$

于是，你期望赢

$$\mathrm{E}(X)=10a\cdot 0.074-a\cdot(1-0.074)=-0.186a=-18.6(元).$$

当只使用一副扑克，可以计算出你每局期望赢-35.29元．

§3.1 数学期望

例3.1.4 在澳门赌场，有很多人在赌廿一点时顺便押对子. 其规则如下：庄家从6副(每副52张)扑克中随机发给你两张. 如果你下注a元，当得到的两张牌是一对时，庄家赔你十倍，否则输掉你的赌注. 如果你下注100元，你在每局中期望赢多少元?

解 用X表示你在一局中的获利，$a=100$. 则

$$P(X=10a)=\frac{13\mathrm{C}_{4\times 6}^{2}}{\mathrm{C}_{52\times 6}^{2}}=0.074,\quad P(X=-a)=1-0.074,$$

于是，你期望赢

$$\mathrm{E}(X)=10a\cdot 0.074-a\cdot(1-0.074)=-0.186a=-18.6(\text{元}).$$

当只使用一副扑克，可以计算出你每局期望赢-35.29元.

§3.1　数学期望

例3.1.5　某个E-mail地址收到相邻的E-mail的时间间隔是独立同分布的随机变量，都服从参数为$\lambda = 0.8\,\text{h}$的指数分布.

(1)　计算两个E-mail之间的平均间隔时间;

(2)　计算从t开始对于下一个E-mail的平均等待时间.

解　(1)用X表示两个E-mail的间隔时间. 根据题意知$X \sim \mathcal{E}(\lambda)$, 有概率密度$f(x) = \lambda e^{-\lambda x}$, $\lambda > 0$. 平均间隔时间是

$$\mathrm{E}(X) = \int_0^{\infty} x f(x)\,\mathrm{d}x = \int_0^{\infty} \lambda x e^{-\lambda x}\,\mathrm{d}x = \frac{1}{0.8} = 1.25(\,\text{h}).$$

(2)　由于指数分布具有无记忆性(参考§3.3 定理3.3.1), 所以从时刻t开始, 需要等待的时间和X同分布, 因而也平均需要等待1.25 h.

§3.1 数学期望

例3.1.5 某个E-mail地址收到相邻的E-mail的时间间隔是独立同分布的随机变量，都服从参数为$\lambda = 0.8\,\mathrm{h}$的指数分布.

(1) 计算两个E-mail之间的平均间隔时间;

(2) 计算从t开始对于下一个E-mail的平均等待时间.

解 (1)用X表示两个E-mail的间隔时间. 根据题意知$X \sim \mathcal{E}(\lambda)$, 有概率密度$f(x) = \lambda e^{-\lambda x}$, $\lambda > 0$. 平均间隔时间是

$$\mathrm{E}(X) = \int_0^{\infty} x f(x)\,\mathrm{d}x = \int_0^{\infty} \lambda x e^{-\lambda x}\,\mathrm{d}x = \frac{1}{0.8} = 1.25(\,\mathrm{h}).$$

(2) 由于指数分布具有无记忆性(参考§3.3 定理3.3.1), 所以从时刻t开始, 需要等待的时间和X同分布, 因而也平均需要等待1.25 h.

§3.1 数学期望

例3.1.5 某个E-mail地址收到相邻的E-mail的时间间隔是独立同分布的随机变量，都服从参数为$\lambda = 0.8\,\mathrm{h}$的指数分布.

(1) 计算两个E-mail之间的平均间隔时间;

(2) 计算从t开始对于下一个E-mail的平均等待时间.

解 (1)用X表示两个E-mail的间隔时间. 根据题意知$X \sim \mathcal{E}(\lambda)$，有概率密度$f(x) = \lambda \mathrm{e}^{-\lambda x}$，$\lambda > 0$. 平均间隔时间是

$$\mathrm{E}(X) = \int_0^{\infty} xf(x)\,\mathrm{d}x = \int_0^{\infty} \lambda x \mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{1}{0.8} = 1.25(\,\mathrm{h}).$$

(2) 由于指数分布具有无记忆性(参考§3.3 定理3.3.1)，所以从时刻t开始，需要等待的时间和X同分布，因而也平均需要等待1.25 h.

§3.1 数学期望

B. 数学期望的统计含义

在例3.1.4中，设甲下注n次，每次下注a元．如果用N_a表示他下注n 次得到的“对子数”，则根据概率的频率定义知道，当$n \to \infty$时，

$$\frac{N_a}{n} \to P(X_1 = 10a) = 0.074,$$
$$\frac{n - N_a}{n} \to P(X_1 = -a) = 1 - 0.074 = 0.926.$$

§3.1 数学期望

B. 数学期望的统计含义

在例3.1.4中，设甲下注n次，每次下注a元. 如果用N_a表示他下注n 次得到的“对子数”，则根据概率的频率定义知道，当$n \to \infty$时，

$$\begin{aligned}\frac{N_a}{n} &\to P(X_1 = 10a) = 0.074,\\ \frac{n - N_a}{n} &\to P(X_1 = -a) = 1 - 0.074 = 0.926.\end{aligned}$$

§3.1 数学期望

用X_i表示他第i次下注的收益，则他独立重复下注n次的平均收益是

$$\begin{aligned}\overline{X}_n &= \frac{X_1+X_2+\cdots+X_n}{n} \\ &= \frac{10aN_a - a(n-N_a)}{n} \\ &= 10a\frac{N_a}{n} - a\frac{n-N_a}{n} \\ &\to 10a\times 0.074 - a\times(1-0.074)\end{aligned}$$

$$= \mathrm{E}(X_1).$$

§5.1 数学期望

以上例子说明，当试验次数增加时，独立重复试验的结果的平均收敛到总体分布的平均. 或者说，当 $n \to \infty$，独立同分布的随机变量的平均$\overline{X}_n$收敛到随机变量的数学期望$\mathrm{E}(X_1) = \mu$.

对于连续型的随机变量，也有相同的结论. 在例1.5中，两个E-mail的间隔时间有数学期望$\mathrm{E}(X) = 1.25$. 用计算机产生10^7个独立同分布的都服从指数分布 $\mathcal{E}(0.8)$ 的随机变量的观测值，利用前 n 个观测值计算的平均数 $\overline{x}_n$ 如下:

n	10	10^2	10^3	10^4	10^5	10^6	10^7
$\overline{x}_n$	1.0439	1.4126	1.3001	1.2531	1.2553	1.2506	1.2501

计算结果支持结论: $\overline{X}_n \to \mathrm{E}(X)$.

§3.2 常用的数学期望

§3.2 常用的数学期望

数学期望的符号E在概率论和统计学中是最常用的符号. 为了简化, 在不引起混淆的情况下, 经常将E后面的括号( )省略. 例如将$\mathrm{E}(X)$写成$\mathrm{E}X$, 将$\mathrm{E}(X^2)$写成$\mathrm{E}X^2$等.

下面是实际中常用分布的数学期望及其计算.

§3.2　常用的数学期望

§3.2　常用的数学期望

数学期望的符号E在概率论和统计学中是最常用的符号. 为了简化, 在不引起混淆的情况下, 经常将E后面的括号(　)省略. 例如将$\mathrm{E}(X)$写成$\mathrm{E}X$, 将$\mathrm{E}(X^2)$写成$\mathrm{E}X^2$等.

下面是实际中常用分布的数学期望及其计算.

§3.2 常用的数学期望

(1) 伯努利分布$\mathcal{B}(1,p)$: 设$X \sim \mathcal{B}(1,p)$, 则

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

又设A是事件, $\mathrm{I}[A]$是A的示性函数, 即

$$\mathrm{I}[A] = \begin{cases} 1, & \text{当}A\text{发生}, \\ 0, & \text{当}A\text{不发生}, \end{cases}$$

则$\mathrm{I}[A]$服从伯努利分布, 且$P(\mathrm{I}[A]=1)=P(A)$. 于是

$$\mathrm{E}\,\mathrm{I}[A] = P(\mathrm{I}[A]=1) = P(A).$$

(2) 二项分布$\mathcal{B}(n,p)$: 设$X \sim \mathcal{B}(n,p)$, 则$\mathrm{E}\,X = np$.

§3.2 常用的数学期望

(1) 伯努利分布$\mathcal{B}(1,p)$: 设$X \sim \mathcal{B}(1,p)$, 则

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

又设A是事件, $\mathrm{I}[A]$是A的示性函数, 即

$$\mathrm{I}[A] = \begin{cases} 1, & \text{当}A\text{发生}, \\ 0, & \text{当}A\text{不发生}, \end{cases}$$

则$\mathrm{I}[A]$服从伯努利分布, 且$P(\mathrm{I}[A]=1)=P(A)$. 于是

$$\mathrm{E}\,\mathrm{I}[A] = P(\mathrm{I}[A]=1) = P(A).$$

(2) 二项分布$\mathcal{B}(n,p)$: 设$X \sim \mathcal{B}(n,p)$, 则$\mathrm{E}\,X = np$.

§3.2 常用的数学期望

(1) 伯努利分布$\mathcal{B}(1,p)$: 设$X \sim \mathcal{B}(1,p)$, 则

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

又设A是事件, $\mathrm{I}[A]$是A的示性函数, 即

$$\mathrm{I}[A] = \begin{cases} 1, & \text{当}A\text{发生}, \\ 0, & \text{当}A\text{不发生}, \end{cases}$$

则$\mathrm{I}[A]$服从伯努利分布, 且$P(\mathrm{I}[A]=1) = P(A)$. 于是

$$\mathrm{E}\,\mathrm{I}[A] = P(\mathrm{I}[A]=1) = P(A).$$

(2) 二项分布$\mathcal{B}(n,p)$: 设$X \sim \mathcal{B}(n,p)$, 则$\mathrm{E}\,X = np$.

§3.2 常用的数学期望

证明 设$q=1-p$，由

$$p_j = P(X=j) = \mathrm{C}_n^j p^j q^{n-j},\ 0 \leqslant j \leqslant n,$$

得到

$$\begin{aligned}
\mathrm{E}X &= \sum_{j=0}^{n} j\mathrm{C}_n^j p^j q^{n-j} \\
&= np\sum_{j=1}^{n} \mathrm{C}_{n-1}^{j-1} p^{j-1} q^{n-j} \quad [\text{用 } j\mathrm{C}_n^j = n\mathrm{C}_{n-1}^{j-1}] \\
&= np\sum_{k=0}^{n-1} \mathrm{C}_{n-1}^{k} p^{k} q^{n-1-k} \qquad [\text{取 } k=j-1] \\
&= np(p+q)^{n-1} = np.
\end{aligned}$$

$\mathrm{E}X=np$说明单次试验成功的概率p越大，则在n次独立重复试验中，平均成功的次数越多.

§3.2 常用的数学期望

证明 设$q=1-p$, 由

$$p_j=P(X=j)=\mathrm{C}_n^j p^j q^{n-j},\ 0\leqslant j\leqslant n,$$

得到

$$\begin{aligned}
\mathrm{E}\,X &= \sum_{j=0}^{n} j\mathrm{C}_n^j p^j q^{n-j} \\
&= np\sum_{j=1}^{n}\mathrm{C}_{n-1}^{j-1}p^{j-1}q^{n-j} \quad [\text{用 } j\mathrm{C}_n^j=n\mathrm{C}_{n-1}^{j-1}] \\
&= np\sum_{k=0}^{n-1}\mathrm{C}_{n-1}^{k}p^{k}q^{n-1-k} \quad [\text{取 } k=j-1] \\
&= np(p+q)^{n-1}=np.
\end{aligned}$$

$\mathrm{E}\,X=np$说明单次试验成功的概率p越大, 则在n次独立重复试验中, 平均成功的次数越多.

§3.2 常用的数学期望

(3) 泊松分布$\mathcal{P}(\lambda)$: 设$X \sim \mathcal{P}(\lambda)$, 则$\mathrm{E}\,X = \lambda$.

证明 由

$$P(X = k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k = 0, 1, \cdots,$$

得到

$$\mathrm{E}\,X = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} = \lambda\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\mathrm{e}^{-\lambda} = \lambda.$$

说明参数λ是泊松分布$\mathcal{P}(\lambda)$的数学期望. 回忆在§3.2的例3.2.1 中, 因为7.5 s 内放射性钋平均放射出3.87个α粒子, 所以当时认为7.5 s 内释放出的粒子数$X \sim \mathcal{P}(3.87)$.

§3.2 常用的数学期望

(3) 泊松分布$\mathcal{P}(\lambda)$: 设$X \sim \mathcal{P}(\lambda)$, 则$\mathrm{E}\, X = \lambda$.

证明 由

$$P(X = k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k = 0, 1, \cdots,$$

得到

$$\mathrm{E}\, X = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}\mathrm{e}^{-\lambda} = \lambda.$$

说明参数λ是泊松分布$\mathcal{P}(\lambda)$的数学期望. 回忆在§3.2的例3.2.1 中, 因为7.5 s 内放射性钋平均放射出3.87个α粒子, 所以当时认为7.5 s 内释放出的粒子数$X \sim \mathcal{P}(3.87)$.

§3.2 常用的数学期望

(4) 几何分布：设X服从参数为p的几何分布，则$\mathrm{E}X = 1/p$.

证明　由

$$P(X=j) = pq^{j-1},\ j = 1, 2, \cdots$$

得到

$$\mathrm{E}X = \sum_{j=1}^{\infty} jpq^{j-1} = p\left(\sum_{j=0}^{\infty} q^j\right)' = p\left(\frac{1}{1-q}\right)' = \frac{1}{p}.$$

结论说明单次试验中的成功概率p越小，首次成功所需要的平均试验次数就越多.

§3.2 常用的数学期望

(4) 几何分布：设X服从参数为p的几何分布，则$\mathrm{E}\,X=1/p$.
证明 由

$$P(X=j)=pq^{j-1},\ j=1,2,\cdots$$

得到

$$\mathrm{E}\,X=\sum_{j=1}^{\infty}jpq^{j-1}=p\Big(\sum_{j=0}^{\infty}q^{j}\Big)'=p\Big(\frac{1}{1-q}\Big)'=\frac{1}{p}.$$

结论说明单次试验中的成功概率p越小，首次成功所需要的平均试验次数就越多.

§3.2 常用的数学期望

(5) 指数分布$\mathcal{E}(\lambda)$: 设$X \sim \mathcal{E}(\lambda)$, 则$\mathrm{E}\,X = 1/\lambda$.

证明 因为X有概率密度

$$f(x) = \lambda \mathrm{e}^{-\lambda x},\ x \geqslant 0,$$

所以

$$\mathrm{E}\,X = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_0^{\infty} x\lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda}.$$

§3.2 常用的数学期望

(5) 指数分布$\mathcal{E}(\lambda)$: 设$X \sim \mathcal{E}(\lambda)$, 则$\mathrm{E}\,X = 1/\lambda$.

证明 因为X有概率密度

$$f(x) = \lambda \mathrm{e}^{-\lambda x},\ x \geqslant 0,$$

所以

$$\mathrm{E}\,X = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_{0}^{\infty} x\lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda}.$$

§3.2 常用的数学期望

定理3.2.1 设X的数学期望有限，概率密度$f(x)$关于c对称：$f(c+x)=f(c-x)$，则$\mathrm{E}X=c$.

证明 这时$g(t)=tf(t+c)$是奇函数：$g(-t)=-g(t)$. 因为$g(t)$在$(-\infty,\infty)$中的积分等于0，所以有

$$\begin{aligned}
\mathrm{E}X &= \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} cf(x)\,\mathrm{d}x + \int_{-\infty}^{\infty}(x-c)f(x-c+c)\,\mathrm{d}x \\
&= c+\int_{-\infty}^{\infty} tf(t+c)\,\mathrm{d}t \\
&= c+\int_{-\infty}^{\infty} g(t)\,\mathrm{d}t \\
&= c.
\end{aligned}$$

§3.2 常用的数学期望

定理3.2.1 设X的数学期望有限，概率密度$f(x)$关于c对称：$f(c+x)=f(c-x)$，则$\mathrm{E}\,X=c$.

证明 这时$g(t)=tf(t+c)$是奇函数：$g(-t)=-g(t)$. 因为$g(t)$在$(-\infty,\infty)$中的积分等于0，所以有

$$\begin{aligned}
\mathrm{E}\,X &= \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} cf(x)\,\mathrm{d}x + \int_{-\infty}^{\infty}(x-c)f(x-c+c)\,\mathrm{d}x \\
&= c+\int_{-\infty}^{\infty} tf(t+c)\,\mathrm{d}t \\
&= c+\int_{-\infty}^{\infty} g(t)\,\mathrm{d}t \\
&= c.
\end{aligned}$$

§3.2 常用的数学期望

定理3.2.1的结论是自然的,因为只要$f(x)$关于c对称, 则c就是曲线$f(x)$和x 轴所夹面积的几何重心的横坐标.

根据概率密度的对称性和定理2.1马上得到如下的推论.

推论3.2.2 正态分布$N(\mu, \sigma^2)$的数学期望是μ, 均匀分布$\mathcal{U}(a, b)$的数学期望是$(a+b)/2$.

§3.2 常用的数学期望

定理3.2.1的结论是自然的,因为只要$f(x)$关于c对称,则c就是曲线$f(x)$和x 轴所夹面积的几何重心的横坐标.

根据概率密度的对称性和定理2.1马上得到如下的推论.

推论3.2.2 正态分布$N(\mu,\sigma^2)$的数学期望是μ,均匀分布$\mathcal{U}(a,b)$的数学期望是$(a+b)/2$.

§3.3 数学期望的计算

§3.3 数学期望的计算

如果$P(X \geqslant 0) = 1$, 则X是非负随机变量. 对于非负的随机变量X, 无论其数学期望$\mathrm{E}X$是否无穷, 都可以直接计算$\mathrm{E}X$.

为了方便计算随机变量函数的数学期望, 再介绍下面的定理.

§3.3 数学期望的计算

§3.3 数学期望的计算

如果$P(X \geqslant 0) = 1$, 则X是非负随机变量. 对于非负的随机变量X, 无论其数学期望$\mathrm{E}\,X$是否无穷, 都可以直接计算$\mathrm{E}\,X$.

为了方便计算随机变量函数的数学期望, 再介绍下面的定理.

§3.3 数学期望的计算

定理3.3.1 设X, Y是随机变量, $\mathrm{E}\, g(X)$, $\mathrm{E}\, h(X,Y)$存在.

(1) 若X有概率密度$f(x)$, 则

$$\mathrm{E}\, g(X) = \int_{-\infty}^{\infty} g(x) f(x)\, \mathrm{d}x. \tag{3.1}$$

(2) 若(X,Y)有联合密度$f(x,y)$, 则

$$\mathrm{E}\, h(X,Y) = \iint_{\mathrm{R}^2} h(x,y) f(x,y)\, \mathrm{d}x\mathrm{d}y. \tag{3.3.2}$$

(3) 若X是非负随机变量, 则

$$\mathrm{E}\, X = \int_{0}^{\infty} P(X > x)\, \mathrm{d}x. \tag{3.3}$$

§3.3 数学期望的计算

定理3.3.1 设X, Y是随机变量, $\mathrm{E}\,g(X)$, $\mathrm{E}\,h(X,Y)$存在.

(1) 若X有概率密度$f(x)$, 则

$$\mathrm{E}\,g(X)=\int_{-\infty}^{\infty} g(x)f(x)\,\mathrm{d}x. \qquad (3.1)$$

(2) 若(X,Y)有联合密度$f(x,y)$, 则

$$\mathrm{E}\,h(X,Y)=\iint_{\mathrm{R}^2} h(x,y)f(x,y)\,\mathrm{d}x\mathrm{d}y. \qquad (3.3.2)$$

(3) 若X是非负随机变量, 则

$$\mathrm{E}\,X=\int_0^{\infty} P(X>x)\,\mathrm{d}x. \qquad (3.3)$$

§3.3 数学期望的计算

使用定理3.3.1计算随机向量函数的数学期望有很多方便，最主要的是不再需要推导随机变量$g(X)$或$g(X,Y)$的概率分布.

例3.3.1 设X在$(0,\pi/2)$上均匀分布，计算$\mathrm{E}(\cos X)$.

解 X有概率密度$f(x)=2/\pi$, $x\in(0,\pi/2)$. 用公式(3.1)得到

$$\mathrm{E}\cos X=\int_{-\infty}^{\infty}f(x)\cos x\,\mathrm{d}x=\frac{2}{\pi}\int_{0}^{\pi/2}\cos x\,\mathrm{d}x=\frac{2}{\pi}.$$

§3.3 数学期望的计算

使用定理3.3.1计算随机向量函数的数学期望有很多方便，最主要的是不再需要推导随机变量$g(X)$或$g(X,Y)$的概率分布.

例3.3.1 设X在$(0,\pi/2)$上均匀分布，计算$\mathrm{E}(\cos X)$.

解 X有概率密度$f(x)=2/\pi,\ x\in(0,\pi/2)$. 用公式(3.1)得到

$$\mathrm{E}\cos X=\int_{-\infty}^{\infty}f(x)\cos x\,\mathrm{d}x=\frac{2}{\pi}\int_0^{\pi/2}\cos x\,\mathrm{d}x=\frac{2}{\pi}.$$

§3.3 数学期望的计算

使用定理3.3.1计算随机向量函数的数学期望有很多方便，最主要的是不再需要推导随机变量$g(X)$或$g(X,Y)$的概率分布.

例3.3.1 设X在$(0,\pi/2)$上均匀分布，计算$\mathrm{E}(\cos X)$.

解 X有概率密度$f(x)=2/\pi,\ x\in(0,\pi/2)$. 用公式(3.1)得到

$$\mathrm{E}\cos X=\int_{-\infty}^{\infty} f(x)\cos x\,\mathrm{d}x=\frac{2}{\pi}\int_{0}^{\pi/2}\cos x\,\mathrm{d}x=\frac{2}{\pi}.$$

§3.3 数学期望的计算

使用定理3.3.1计算随机向量函数的数学期望有很多方便，最主要的是不再需要推导随机变量$g(X)$或$g(X,Y)$的概率分布.

例3.3.1 设X在$(0,\pi/2)$上均匀分布，计算$\mathrm{E}(\cos X)$.

解 X有概率密度$f(x)=2/\pi,\ x\in(0,\pi/2)$. 用公式(3.1)得到

$$\mathrm{E}\cos X=\int_{-\infty}^{\infty}f(x)\cos x\,\mathrm{d}x=\frac{2}{\pi}\int_{0}^{\pi/2}\cos x\,\mathrm{d}x=\frac{2}{\pi}.$$

§3.3 数学期望的计算

例3.3.2 设X, Y 独立, 都服从标准正态分布, 计算$\mathrm{E}\,(X^2+Y^2)$.

解 (X,Y) 有联合密度

$$f(x,y)=\frac{1}{2\pi}\exp\Big(-\frac{x^2+y^2}{2}\Big).$$

用公式(3.2), 且在积分中采用变换$x=r\cos\theta$, $y=r\sin\theta$, 得到

$$\begin{aligned}\mathrm{E}\,(X^2+Y^2) &= \iint_{\mathrm{R}^2}(x^2+y^2)f(x,y)\,\mathrm{d}x\mathrm{d}y\\ &= \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta\int_0^{\infty}r^3\exp(-r^2/2)\,\mathrm{d}r\\ &= \int_0^{\infty}r^3\exp(-r^2/2)\,\mathrm{d}r \qquad [\text{取}\,t=r^2/2]\\ &= 2\int_0^{\infty}t\exp(-t)\,\mathrm{d}t\\ &= 2\Gamma(2)=2\cdot(2-1)!=2.\end{aligned}$$

§3.3 数学期望的计算

例3.3.2 设X, Y 独立, 都服从标准正态分布, 计算$\mathrm{E}\,(X^2+Y^2)$.

解 (X,Y) 有联合密度

$$f(x,y)=\frac{1}{2\pi}\exp\Big(-\frac{x^2+y^2}{2}\Big).$$

用公式(3.2), 且在积分中采用变换$x=r\cos\theta$, $y=r\sin\theta$, 得到

$$\begin{aligned}
\mathrm{E}\,(X^2+Y^2) &= \iint_{\mathrm{R}^2}(x^2+y^2)f(x,y)\,\mathrm{d}x\mathrm{d}y\\
&= \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta\int_0^{\infty} r^3\exp(-r^2/2)\,\mathrm{d}r\\
&= \int_0^{\infty} r^3\exp(-r^2/2)\,\mathrm{d}r \qquad [\text{取}\,t=r^2/2]\\
&= 2\int_0^{\infty} t\exp(-t)\,\mathrm{d}t\\
&= 2\Gamma(2)=2\cdot(2-1)!=2.
\end{aligned}$$

§3.3 数学期望的计算

例3.3.2 设X, Y 独立, 都服从标准正态分布, 计算$\mathrm{E}\,(X^2+Y^2)$.

解 (X, Y) 有联合密度

$$f(x,y)=\frac{1}{2\pi}\exp\Big(-\frac{x^2+y^2}{2}\Big).$$

用公式(3.2), 且在积分中采用变换$x=r\cos\theta$, $y=r\sin\theta$, 得到

$$\begin{aligned}
\mathrm{E}\,(X^2+Y^2) &= \iint_{\mathrm{R}^2}(x^2+y^2)f(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta\int_0^{\infty} r^3\exp(-r^2/2)\,\mathrm{d}r \\
&= \int_0^{\infty} r^3\exp(-r^2/2)\,\mathrm{d}r \qquad [\text{取}\, t=r^2/2] \\
&= 2\int_0^{\infty} t\exp(-t)\,\mathrm{d}t \\
&= 2\Gamma(2)=2\cdot(2-1)!=2.
\end{aligned}$$

§3.3 数学期望的计算

例3.3.3 秘书长的3台传真机独立工作. 第j台传真机对下一个到达的传真的等待时间服从参数为λ_j的指数分布. 计算该秘书长对第一个到达的传真的平均等待时间.

解 用X_j表示第j台传真机对下一个传真的等待时间, 则X_1, X_2, X_3 相互独立, $X_j \sim \mathcal{E}(\lambda_j)$. $Y = \min(X_1, X_2, X_3)$是对第一个传真的等待时间. 容易计算

$$P(X_j > y) = \int_y^{\infty} \lambda_j \mathrm{e}^{-\lambda_j x} \,\mathrm{d}x = \mathrm{e}^{-\lambda_j y}.$$

$$\begin{aligned} P(Y > y) &= P(X_1 > y, X_2 > y, X_3 > y) \\ &= P(X_1 > y)P(X_2 > y)P(X_3 > y) \\ &= \mathrm{e}^{-(\lambda_1+\lambda_2+\lambda_3)y}. \end{aligned}$$

§3.3 数学期望的计算

例3.3.3 秘书长的3台传真机独立工作. 第j台传真机对下一个到达的传真的等待时间服从参数为λ_j的指数分布. 计算该秘书长对第一个到达的传真的平均等待时间.

解 用X_j表示第j台传真机对下一个传真的等待时间, 则X_1, X_2, X_3 相互独立, $X_j \sim \mathcal{E}(\lambda_j)$. $Y = \min(X_1, X_2, X_3)$是对第一个传真的等待时间. 容易计算

$$P(X_j > y) = \int_y^{\infty} \lambda_j \mathrm{e}^{-\lambda_j x}\,\mathrm{d}x = \mathrm{e}^{-\lambda_j y}.$$

$$\begin{aligned}
P(Y > y) &= P(X_1 > y, X_2 > y, X_3 > y) \\
&= P(X_1 > y)P(X_2 > y)P(X_3 > y) \\
&= \mathrm{e}^{-(\lambda_1+\lambda_2+\lambda_3)y}.
\end{aligned}$$

§3.3 数学期望的计算

例3.3.3 秘书长的3台传真机独立工作. 第j台传真机对下一个到达的传真的等待时间服从参数为λ_j的指数分布. 计算该秘书长对第一个到达的传真的平均等待时间.

解 用X_j表示第j台传真机对下一个传真的等待时间, 则X_1, X_2, X_3 相互独立, $X_j \sim \mathcal{E}(\lambda_j)$. $Y = \min(X_1, X_2, X_3)$是对第一个传真的等待时间. 容易计算

$$P(X_j > y) = \int_y^{\infty} \lambda_j \mathrm{e}^{-\lambda_j x}\, \mathrm{d}x = \mathrm{e}^{-\lambda_j y}.$$

$$\begin{aligned} P(Y > y) &= P(X_1 > y, X_2 > y, X_3 > y) \\ &= P(X_1 > y)P(X_2 > y)P(X_3 > y) \\ &= \mathrm{e}^{-(\lambda_1+\lambda_2+\lambda_3)y}. \end{aligned}$$

§3.3　数学期望的计算

再用定理3.3.1(3)得到

$$\mathrm{E}\,Y = \int_0^{\infty} \mathrm{e}^{-(\lambda_1+\lambda_2+\lambda_3)y}\,\mathrm{d}y = \frac{1}{\lambda_1+\lambda_2+\lambda_3}.$$

由于指数分布具有无记忆性，所以例3.3中的等待时间可以是从任何时刻开始的等待时间.

对于离散型随机变量和随机向量也有类似定理3.1的结论，这就是下面的定理3.2.

§3.3 数学期望的计算

再用定理3.3.1(3)得到

$$\mathrm{E}\,Y = \int_0^{\infty} \mathrm{e}^{-(\lambda_1+\lambda_2+\lambda_3)y}\,\mathrm{d}y = \frac{1}{\lambda_1+\lambda_2+\lambda_3}.$$

由于指数分布具有无记忆性，所以例3.3中的等待时间可以是从任何时刻开始的等待时间.

对于离散型随机变量和随机向量也有类似定理3.1的结论，这就是下面的定理3.2.

§3.3 数学期望的计算

定理3.3.2 设X, Y是离散型随机变量, $\mathrm{E}\, g(X)$, $\mathrm{E}\, h(X,Y)$存在.

(1) 若X有离散分布$p_j=P(X=x_j)$, $j\geqslant 1$, 则

$$\mathrm{E}\, g(X)=\sum_{j=1}^{\infty} g(x_j)p_j.$$

(2) 若(X,Y)有离散分布$p_{ij}=P(X=x_i, Y=y_j)$, $i,j\geqslant 1$, 则

$$\mathrm{E}\, h(X,Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} h(x_i,y_j)p_{ij}.$$

§3.3 数学期望的计算

定理3.3.2 设X, Y是离散型随机变量, $\mathrm{E}\, g(X)$, $\mathrm{E}\, h(X, Y)$存在.

(1) 若X有离散分布$p_j = P(X = x_j)$, $j \geqslant 1$, 则

$$\mathrm{E}\, g(X) = \sum_{j=1}^{\infty} g(x_j) p_j.$$

(2) 若(X, Y)有离散分布$p_{ij} = P(X = x_i, Y = y_j)$, $i, j \geqslant 1$, 则

$$\mathrm{E}\, h(X, Y) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h(x_i, y_j) p_{ij}.$$

§3.3 数学期望的计算

例3.3.4 设X服从二项分布$\mathcal{B}(n,p)$, 计算$\mathrm{E}[X(X-1)]$.

解 从$P(X=j)=\mathrm{C}_n^j p^j q^{n-j}$知道

$$
\begin{aligned}
\mathrm{E}[X(X-1)] &= \sum_{j=0}^{n} j(j-1)\mathrm{C}_n^j p^j q^{n-j} \\
&= p^2\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sum_{j=0}^{n}\mathrm{C}_n^j x^j q^{n-j}\right)\Big|_{x=p} \\
&= p^2\frac{\mathrm{d}^2}{\mathrm{d}x^2}(x+q)^n\Big|_{x=p} \\
&= n(n-1)p^2.
\end{aligned}
$$

§3.3 数学期望的计算

例3.3.4 设X服从二项分布$\mathcal{B}(n,p)$, 计算$\mathrm{E}[X(X-1)]$.

解 从$P(X=j)=\mathrm{C}_n^j p^j q^{n-j}$知道

$$\begin{aligned}\mathrm{E}[X(X-1)] &= \sum_{j=0}^{n} j(j-1)\mathrm{C}_n^j p^j q^{n-j} \\ &= p^2\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sum_{j=0}^{n}\mathrm{C}_n^j x^j q^{n-j}\right)\Big|_{x=p} \\ &= p^2\frac{\mathrm{d}^2}{\mathrm{d}x^2}(x+q)^n\Big|_{x=p} \\ &= n(n-1)p^2.\end{aligned}$$

§3.3 数学期望的计算

例3.3.4 设X服从二项分布$\mathcal{B}(n,p)$，计算$\mathrm{E}\,[X(X-1)]$.

解 从$P(X=j)=\mathrm{C}_n^j p^j q^{n-j}$知道

$$
\begin{aligned}
\mathrm{E}[X(X-1)] &= \sum_{j=0}^{n} j(j-1)\mathrm{C}_n^j p^j q^{n-j} \\
&= p^2\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sum_{j=0}^{n}\mathrm{C}_n^j x^j q^{n-j}\right)\Big|_{x=p} \\
&= p^2\frac{\mathrm{d}^2}{\mathrm{d}x^2}(x+q)^n\Big|_{x=p} \\
&= n(n-1)p^2.
\end{aligned}
$$

§3.4 数学期望的性质

§3.4 数学期望的性质

根据定理3.3.1和定理3.3.2,

$$\mathrm{E}\,|X| = \begin{cases} \sum\limits_{j=1}^{\infty} |x_j| P(X = x_j), & \text{当} \sum\limits_{j=1}^{\infty} P(X = x_j) = 1, \\ \int_{-\infty}^{\infty} |x| f(x)\,\mathrm{d}x, & \text{当} X \text{有概率密度} f(x). \end{cases}$$

于是, $\mathrm{E}\,X$存在的充分必要条件是$\mathrm{E}\,|X| < \infty$.

§3.4 数学期望的性质

§3.4 数学期望的性质

根据定理3.3.1和定理3.3.2,

$$\mathrm{E}\,|X| = \begin{cases} \sum_{j=1}^{\infty} |x_j| P(X = x_j), & \text{当} \sum_{j=1}^{\infty} P(X = x_j) = 1, \\ \int_{-\infty}^{\infty} |x| f(x)\,\mathrm{d}x, & \text{当}X\text{有概率密度}f(x). \end{cases}$$

于是, $\mathrm{E}\,X$存在的充分必要条件是$\mathrm{E}\,|X| < \infty$.

§3.4 数学期望的性质

定理3.4.1 设$\mathrm{E}\,|X_j| < \infty$ $(1 \leqslant j \leqslant n)$, $c_0, c_1, \cdots, c_n$是常数, 则有以下结果:

(1) 线性组合$Y = c_0 + c_1X_1 + c_2X_2 + \cdots + c_nX_n$的数学期望存在, 而且

$$\begin{aligned} & \mathrm{E}\,(c_0 + c_1X_1 + c_2X_2 + \cdots + c_nX_n) \\ = \; & c_0 + c_1\mathrm{E}\,X_1 + c_2\mathrm{E}\,X_2 + \cdots + c_n\mathrm{E}\,X_n; \end{aligned} \tag{4.1}$$

(2) 如果$X_1, X_2, \cdots, X_n$相互独立, 则乘积$Z = X_1X_2\cdots X_n$的数学期望存在, 并且

$$\mathrm{E}\,(X_1X_2\cdots X_n) = \mathrm{E}\,X_1\mathrm{E}\,X_2\cdots\mathrm{E}\,X_n;$$

§3.4 数学期望的性质

定理3.4.1 设$\mathrm{E}\,|X_j| < \infty$ $(1 \leqslant j \leqslant n)$, $c_0, c_1, \cdots, c_n$是常数, 则有以下结果:

(1) 线性组合$Y = c_0 + c_1X_1 + c_2X_2 + \cdots + c_nX_n$的数学期望存在, 而且

$$\begin{aligned} & \mathrm{E}\,(c_0 + c_1X_1 + c_2X_2 + \cdots + c_nX_n) \\ = \quad & c_0 + c_1\mathrm{E}\,X_1 + c_2\mathrm{E}\,X_2 + \cdots + c_n\mathrm{E}\,X_n; \end{aligned} \tag{4.1}$$

(2) 如果$X_1, X_2, \cdots, X_n$相互独立, 则乘积$Z = X_1X_2\cdots X_n$的数学期望存在, 并且

$$\mathrm{E}\,(X_1X_2\cdots X_n) = \mathrm{E}\,X_1\mathrm{E}\,X_2\cdots\mathrm{E}\,X_n;$$

§3.4 数学期望的性质

(3) 如果$P(X_1 \leqslant X_2) = 1$, 则$\mathrm{E}\, X_1 \leqslant \mathrm{E}\, X_2$.

性质(1)说明对随机变量求数学期望的运算是线性运算; 性质(2)说明相互独立的随机变量积的数学期望等于数学期望的积; 性质(3)说明如果$X_1 \leqslant X_2$, 则对X_1的期望值应小于等于对X_2的期望值.

§3.4 数学期望的性质

(3) 如果$P(X_1 \leqslant X_2) = 1$, 则$\mathrm{E}\,X_1 \leqslant \mathrm{E}\,X_2$.

性质(1)说明对随机变量求数学期望的运算是线性运算; 性质(2)说明相互独立的随机变量积的数学期望等于数学期望的积; 性质(3)说明如果$X_1 \leqslant X_2$, 则对X_1的期望值应小于等于对X_2的期望值.

§3.4 数学期望的性质

例3.4.1 设$X \sim N(0,1)$, 则$\mathrm{E}\,X^2 = 1$.

证明 取随机变量Y和X独立同分布, 则$\mathrm{E}\,Y^2 = \mathrm{E}\,X^2$. 从例3.2的结论知道$\mathrm{E}\,(X^2+Y^2) = 2$, 于是有

$$\mathrm{E}\,X^2 = (\mathrm{E}\,X^2 + \mathrm{E}\,Y^2)/2 = \mathrm{E}\,(X^2+Y^2)/2 = 2/2 = 1.$$

X的数学期望是指对X的期望值, 它也是X的平均取值. 看下面的例子.

§3.4 数学期望的性质

例3.4.1 设$X \sim N(0,1)$, 则$\mathrm{E}\,X^2 = 1$.

证明 取随机变量Y和X独立同分布, 则$\mathrm{E}\,Y^2 = \mathrm{E}\,X^2$. 从例3.2的结论知道$\mathrm{E}\,(X^2+Y^2) = 2$, 于是有

$$\mathrm{E}\,X^2 = (\mathrm{E}\,X^2 + \mathrm{E}\,Y^2)/2 = \mathrm{E}\,(X^2+Y^2)/2 = 2/2 = 1.$$

X的数学期望是指对X的期望值, 它也是X的平均取值. 看下面的例子.

§3.4 数学期望的性质

例3.4.1 设$X \sim N(0,1)$, 则$\mathrm{E}\,X^2 = 1$.

证明 取随机变量Y和X独立同分布, 则$\mathrm{E}\,Y^2 = \mathrm{E}\,X^2$. 从例3.2的结论知道$\mathrm{E}\,(X^2 + Y^2) = 2$, 于是有

$$\mathrm{E}\,X^2 = (\mathrm{E}\,X^2 + \mathrm{E}\,Y^2)/2 = \mathrm{E}\,(X^2 + Y^2)/2 = 2/2 = 1.$$

X的数学期望是指对X的期望值, 它也是X的平均取值. 看下面的例子.

§3.4 数学期望的性质

例3.4.2(二项分布$\mathcal{B}(n,p)$) 设单次试验成功的概率是p, 问n次独立重复试验中, 期望有几次成功?

解 引入

$$X_i = \begin{cases} 1, & \text{第}i\text{次试验成功}, \\ 0, & \text{第}i\text{次试验不成功}, \end{cases}$$

则$\mathrm{E}\,X_i = p$. $X = X_1 + X_2 + \cdots + X_n$是$n$次试验中的成功次数, 服从二项分布$\mathcal{B}(n,p)$. 期望的成功次数是

$$\mathrm{E}\,X = \mathrm{E}\,X_1 + \mathrm{E}\,X_2 + \cdots + \mathrm{E}\,X_n = np.$$

§3.4 数学期望的性质

例3.4.2(二项分布$\mathcal{B}(n,p)$) 设单次试验成功的概率是p, 问n次独立重复试验中, 期望有几次成功?

解 引入

$$X_i = \begin{cases} 1, & \text{第}i\text{次试验成功}, \\ 0, & \text{第}i\text{次试验不成功}, \end{cases}$$

则$\mathrm{E}\,X_i = p$. $X = X_1 + X_2 + \cdots + X_n$是$n$次试验中的成功次数, 服从二项分布$\mathcal{B}(n,p)$. 期望的成功次数是

$$\mathrm{E}\,X = \mathrm{E}\,X_1 + \mathrm{E}\,X_2 + \cdots + \mathrm{E}\,X_n = np.$$

§3.4 数学期望的性质

例3.4.2(二项分布$\mathcal{B}(n,p)$) 设单次试验成功的概率是p，问n次独立重复试验中，期望有几次成功？

解 引入

$$X_i=\begin{cases}1, & \text{第}i\text{次试验成功},\\ 0, & \text{第}i\text{次试验不成功},\end{cases}$$

则$\mathrm{E}\,X_i=p$. $X=X_1+X_2+\cdots+X_n$是n次试验中的成功次数，服从二项分布$\mathcal{B}(n,p)$. 期望的成功次数是

$$\mathrm{E}\,X=\mathrm{E}\,X_1+\mathrm{E}\,X_2+\cdots+\mathrm{E}\,X_n=np.$$

§3.4 数学期望的性质

例3.4.3(超几何分布$H(N,M,n)$) N件产品中有M件正品，从中任取n件，期望有几件正品?

解 定义随机变量

$$X_i=\begin{cases}1, & \text{第}i\text{次取得正品},\\ 0, & \text{第}i\text{次取得次品},\end{cases}$$

则无论是否有放回地抽取，总有$\mathrm{E}\,X_i=M/N$(参考抽签问题)。无放回抽取时，抽到的正品数$Y=(X_1+X_2+\cdots+X_n)$服从超几何分布$H(N,M,n)$。期望的正品数是

$$\mathrm{E}\,Y=\mathrm{E}\,X_1+\mathrm{E}\,X_2+\cdots+\mathrm{E}\,X_n=nM/N.$$

本例中，如果有放回地抽取，则$Y\sim\mathcal{B}(n,M/N)$。如果无放回地抽取，则$Y\sim H(N,M,n)$。无论是否有放回地抽取，期望得到的正品数都是n倍的正品率。

§3.4 数学期望的性质

例3.4.3(超几何分布$H(N, M, n)$) N件产品中有M件正品，从中任取n 件，期望有几件正品？

解 定义随机变量

$$X_i = \begin{cases} 1, & \text{第}i\text{次取得正品}, \\ 0, & \text{第}i\text{次取得次品}, \end{cases}$$

则无论是否有放回地抽取，总有$\mathrm{E}\,X_i = M/N$(参考抽签问题)。无放回抽取时，抽到的正品数$Y = (X_1 + X_2 + \cdots + X_n)$服从超几何分布$H(N, M, n)$。期望的正品数是

$$\mathrm{E}\,Y = \mathrm{E}\,X_1 + \mathrm{E}\,X_2 + \cdots + \mathrm{E}\,X_n = nM/N.$$

本例中，如果有放回地抽取，则$Y \sim \mathcal{B}(n, M/N)$。如果无放回地抽取，则$Y \sim H(N, M, n)$。无论是否有放回地抽取，期望得到的正品数都是$n$倍的正品率。

§3.4 数学期望的性质

例3.4.3(超几何分布$H(N,M,n)$) N件产品中有M件正品，从中任取n 件，期望有几件正品？

解 定义随机变量

$$X_i = \begin{cases} 1, & \text{第}i\text{次取得正品}, \\ 0, & \text{第}i\text{次取得次品}, \end{cases}$$

则无论是否有放回地抽取，总有$\mathrm{E}\,X_i = M/N$(参考抽签问题)。无放回抽取时，抽到的正品数$Y = (X_1 + X_2 + \cdots + X_n)$服从超几何分布$H(N,M,n)$。期望的正品数是

$$\mathrm{E}\,Y = \mathrm{E}\,X_1 + \mathrm{E}\,X_2 + \cdots + \mathrm{E}\,X_n = nM/N.$$

本例中，如果有放回地抽取，则$Y \sim \mathcal{B}(n, M/N)$。如果无放回地抽取，则$Y \sim H(N,M,n)$。无论是否有放回地抽取，期望得到的正品数都是$n$倍的正品率。

§3.4 数学期望的性质

例3.4.4 将n个不同的信笺随机放入n个写好地址的信封，期望有几封能正确搭配？

解 定义随机变量

$$X_i = \begin{cases} 1, & \text{第}\,i\,\text{封信正确搭配}, \\ 0, & \text{第}\,i\,\text{封信没有正确搭配}, \end{cases}$$

则$\mathrm{E}\,X_i = P(X_i = 1) = 1/n$. 因为$Y = X_1 + X_2 + \cdots + X_n$是正确搭配的个数，所以平均正确搭配的个数是

$$\mathrm{E}\,Y = \mathrm{E}\,X_1 + \mathrm{E}\,X_2 + \cdots + \mathrm{E}\,X_n = n/n = 1.$$

本例说明无论有多少个信封，平均只有一封信能正确搭配.

§3.4 数学期望的性质

例3.4.4 将n个不同的信笺随机放入n个写好地址的信封，期望有几封能正确搭配？

解 定义随机变量

$$X_i = \begin{cases} 1, & 第i封信正确搭配, \\ 0, & 第i封信没有正确搭配, \end{cases}$$

则$\mathrm{E}\,X_i = P(X_i = 1) = 1/n$. 因为$Y = X_1 + X_2 + \cdots + X_n$是正确搭配的个数，所以平均正确搭配的个数是

$$\mathrm{E}\,Y = \mathrm{E}\,X_1 + \mathrm{E}\,X_2 + \cdots + \mathrm{E}\,X_n = n/n = 1.$$

本例说明无论有多少个信封，平均只有一封信能正确搭配.

§3.4 数学期望的性质

例3.4.4 将n个不同的信笺随机放入n个写好地址的信封，期望有几封能正确搭配？

解 定义随机变量

$$X_i = \begin{cases} 1, & \text{第}i\text{封信正确搭配}, \\ 0, & \text{第}i\text{封信没有正确搭配}, \end{cases}$$

则$\mathrm{E}\,X_i = P(X_i = 1) = 1/n$. 因为$Y = X_1 + X_2 + \cdots + X_n$是正确搭配的个数，所以平均正确搭配的个数是

$$\mathrm{E}\,Y = \mathrm{E}\,X_1 + \mathrm{E}\,X_2 + \cdots + \mathrm{E}\,X_n = n/n = 1.$$

本例说明无论有多少个信封，平均只有一封信能正确搭配.

§3.4 数学期望的性质

例3.4.5 设商店每销售一袋大米获利a元，每库存一袋大米损失b元，假设大米的销量Y服从指数分布$\mathcal{E}(\lambda)$. 问库存多少袋大米才能获得最大的平均利润.

解 库存量是x时，利润是

$$Q(x,Y)=\begin{cases} aY-b(x-Y), & Y<x, \\ ax, & Y\geqslant x. \end{cases}$$

用$\mathrm{I}[Y<x]$表示事件$\{Y<x\}$的示性函数，用$\mathrm{I}[Y\geqslant x]$表示$\{Y\geqslant x\}$的示性函数，则可以将$Q(x,Y)$写成

$$Q(x,Y)=\big[aY-b(x-Y)\big]\mathrm{I}[Y<x]+ax\mathrm{I}[Y\geqslant x].$$

§3.4 数学期望的性质

例3.4.5 设商店每销售一袋大米获利a元，每库存一袋大米损失b元，假设大米的销量Y服从指数分布$\mathcal{E}(\lambda)$. 问库存多少袋大米才能获得最大的平均利润.

解 库存量是x时，利润是

$$Q(x,Y)=\begin{cases} aY-b(x-Y), & Y<x,\\ ax, & Y\geq x. \end{cases}$$

用$\mathrm{I}[Y<x]$表示事件$\{Y<x\}$的示性函数，用$\mathrm{I}[Y\geq x]$表示$\{Y\geq x\}$的示性函数，则可以将$Q(x,Y)$写成

$$Q(x,Y)=\big[aY-b(x-Y)\big]\mathrm{I}[Y<x]+ax\mathrm{I}[Y\geq x].$$

§3.4 数学期望的性质

例3.4.5 设商店每销售一袋大米获利a元，每库存一袋大米损失b元，假设大米的销量Y服从指数分布$\mathcal{E}(\lambda)$. 问库存多少袋大米才能获得最大的平均利润.

解 库存量是x时，利润是

$$Q(x, Y) = \begin{cases} aY - b(x - Y), & Y < x, \\ ax, & Y \geqslant x. \end{cases}$$

用$\mathrm{I}[Y < x]$表示事件$\{Y < x\}$的示性函数，用$\mathrm{I}[Y \geqslant x]$表示$\{Y \geqslant x\}$的示性函数，则可以将$Q(x, Y)$写成

$$Q(x, Y) = \big[aY - b(x - Y)\big]\mathrm{I}[Y < x] + ax\mathrm{I}[Y \geqslant x].$$

§3.4 数学期望的性质

Y有概率密度$f_Y(y) = \lambda \mathrm{e}^{-\lambda y}, y > 0$. 所以平均利润是

$$\begin{aligned} q(x) = \quad & \mathrm{E}\, Q(x, Y) \\ = \quad & \int_{-\infty}^{\infty} Q(x, y) f_Y(y)\, \mathrm{d}y \\ = \quad & \mathrm{E}\,\{[aY - b(x - Y)]\mathrm{I}[Y < x] + ax\mathrm{I}[Y \geqslant x]\} \\ = \quad & \int_0^x [ay - b(x - y)]\, f_Y(y)\, \mathrm{d}y + ax \int_x^{\infty} f_Y(y)\, \mathrm{d}y \\ = \quad & (a + b)(1 - \mathrm{e}^{-\lambda x})/\lambda - bx. \end{aligned}$$

$q(x)$ 的最大值点是所要的库存数.

§3.4 数学期望的性质

Y有概率密度$f_Y(y)=\lambda \mathrm{e}^{-\lambda y}, y>0$. 所以平均利润是

$$
\begin{aligned}
q(x) = \quad & \mathrm{E}\, Q(x,Y) \\
= \quad & \int_{-\infty}^{\infty} Q(x,y) f_Y(y)\,\mathrm{d}y \\
= \quad & \mathrm{E}\,\{[aY - b(x-Y)]\mathrm{I}[Y<x] + ax\mathrm{I}[Y\geqslant x]\} \\
= \quad & \int_0^x \big[ay - b(x-y)\big] f_Y(y)\,\mathrm{d}y + ax\int_x^{\infty} f_Y(y)\,\mathrm{d}y \\
= \quad & (a+b)(1-\mathrm{e}^{-\lambda x})/\lambda - bx.
\end{aligned}
$$

$q(x)$ 的最大值点是所要的库存数.

§3.4 数学期望的性质

由

$$q'(x) = (a+b)\mathrm{e}^{-\lambda x} - b = 0$$

得到$q(x)$的唯一极值点$x = \lambda^{-1}\ln[(a+b)/b]$. 由

$$q''(x) = -(a+b)\lambda\mathrm{e}^{-\lambda x} < 0$$

知道$q(x)$是上凸函数. 所以, $x = \lambda^{-1}\ln[(a+b)/b]$是$q(x)$的唯一最大值点. 于是, 库存$\lambda^{-1}\ln[(a+b)/b]$袋大米可以获得最大平均利润.

§3.4 数学期望的性质

由

$$q'(x) = (a+b)\mathrm{e}^{-\lambda x} - b = 0$$

得到$q(x)$的唯一极值点$x = \lambda^{-1}\ln[(a+b)/b]$. 由

$$q''(x) = -(a+b)\lambda\mathrm{e}^{-\lambda x} < 0$$

知道$q(x)$是上凸函数. 所以, $x = \lambda^{-1}\ln[(a+b)/b]$是$q(x)$的唯一最大值点. 于是, 库存$\lambda^{-1}\ln[(a+b)/b]$袋大米可以获得最大平均利润.

§3.4 数学期望的性质

定理3.4.2 $\mathrm{E}|X| = 0$的充分必要条件是$P(X = 0) = 1$.

如果$P(X = 0) = 1$, 则称$X = 0$ 以概率1发生, 记做$X = 0$ a.s.. 完全类似地, 我们把$P(X \leqslant Y) = 1$记做$X \leqslant Y$ a.s.. 当$P(A) = 1$, 我们称A 以概率1 发生. 以概率1发生又称作几乎处处或几乎必然(almost surely)发生.

§3.4 数学期望的性质

定理3.4.2 $\mathrm{E}\,|X| = 0$的充分必要条件是$P(X = 0) = 1$.

如果$P(X = 0) = 1$, 则称$X = 0$ 以概率1发生, 记做$X = 0$ a.s.. 完全类似地, 我们把$P(X \leqslant Y) = 1$记做$X \leqslant Y$ a.s.. 当$P(A) = 1$, 我们称A 以概率1 发生. 以概率1发生又称作几乎处处或几乎必然(almost surely)发生.

§3.5 条件数学期望

在随机向量(X, Y) 中，给定随机变量 X 的取值 $X = x$ 时Y 的条件期望. 例如，人的基本健康状况可以用身高(H)、体重(W)组成的随机向量(H,W) 来表达，给定一个人的身高$H = h$ 时，求得的平均体重就是条件数学期望, 简称条件期望. 通常用 $\mathrm{E}(Y|X = x)$ 表示，也可以简化记为 $\mathrm{E}(Y|x)$.

定义3.5.1 (随机变量的条件期望)

- 对离散型随机变量X 和Y, X 取值于$\{x_1, x_2, \cdots, x_n\}$, Y 取值于 $\{y_1, y_2, \cdots, y_m\}$, 则

$$\mathrm{E}(X|Y = y_k) = \sum_{i=1}^{n} \mathrm{P}(X = x_i|Y = y_k).$$

- 对连续型随机变量$(X, Y) \sim f(x, y)$, 记给定$X = x$ 时随机变量Y 的条件密度函数为 $f_{Y|X}(y|x)$, 设$\mathrm{E}(|Y| < \infty$, 则

$$\mathrm{E}(Y|X = x) = \int y f_{Y|X}(y|x) dy.$$

§3.5 条件数学期望

在随机向量(X,Y) 中，给定随机变量 X 的取值 $X=x$ 时Y 的条件期望. 例如，人的基本健康状况可以用身高(H)、体重(W)组成的随机向量(H,W) 来表达，给定一个人的身高$H=h$ 时，求得的平均体重就是条件数学期望, 简称条件期望. 通常用 $\mathrm{E}(Y|X=x)$ 表示，也可以简化记为 $\mathrm{E}(Y|x)$.

定义3.5.1 (随机变量的条件期望)

- 对离散型随机变量X 和Y, X 取值于$\{x_1,x_2,\cdots,x_n\}$, Y 取值于 $\{y_1,y_2,\cdots,y_m\}$, 则

$$\mathrm{E}(X|Y=y_k)=\sum_{i=1}^{n}\mathrm{P}(X=x_i|Y=y_k).$$

- 对连续型随机变量$(X,Y)\sim f(x,y)$, 记给定$X=x$ 时随机变量Y 的条件密度函数为 $f_{Y|X}(y|x)$, 设$\mathrm{E}(|Y|<\infty$, 则

$$\mathrm{E}(Y|X=x)=\int yf_{Y|X}(y|x)dy.$$

§3.5　条件数学期望

例3.5.1　设$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，给定$X=x$ 时随机变量Y 的条件分布仍是正态分布，即

$$Y|x\sim N(\mu_2+\rho\sigma_2\sigma_1^{-1}(x-\mu_1),(1-\rho^2)\sigma_2^2),$$

从而条件期望为

$$\mathrm{E}(Y|x)=\mu_2+\mu_2+\rho\sigma_2\sigma_1^{-1}(x-\mu_1).$$

这是x 的线性函数，当$\rho>0$ 时，$\mathrm{E}(Y|x)$ 随X 增加而增加，即Y 的均值有随X 的增加而增加的趋势.

§3.5 条件数学期望

例3.5.2 (巴格达窃贼问题) 一窃贼被关在有3 扇门的地牢里，其中1 号门通向自由. 出这扇门走3 h 便可以回到地面；2 号门通向另一个地道，走5 h 将返回到地牢；3 号门通向更长的地道，走7 h 也回到地牢. 若窃贼每次选择3 扇门的可能性总相同，求他为获得自由而奔走的平均时间.

解: 设这个窃贼需要走Xh 才能到达地面，并设Y 代表他每次对3 扇门的选择情况，Y 各以$\frac{1}{3}$ 的概率取值1,2,3. 则

$$\mathrm{E}(X) = \mathrm{E}[E(X|Y)] = \sum_{i=1}^{3} \mathrm{E}(X|Y=i)\mathrm{P}(Y=i),$$

注意到$\mathrm{E}(X|Y=1)=3$, $\mathrm{E}(X|Y=2)=5+\mathrm{E}(X)$, $\mathrm{E}(X|Y=3)=7+\mathrm{E}(X)$, 所以

$$\mathrm{E}(X) = \frac{1}{3}[3+5+\mathrm{E}(X)+7+E(X)].$$

即得到$\mathrm{E}(X)=15$.

§3.5 条件数学期望

例3.5.2 (巴格达窃贼问题) 一窃贼被关在有3 扇门的地牢里，其中1 号门通向自由. 出这扇门走3 h 便可以回到地面；2 号门通向另一个地道，走5 h 将返回到地牢；3 号门通向更长的地道，走7 h 也回到地牢. 若窃贼每次选择3 扇门的可能性总相同，求他为获得自由而奔走的平均时间.

解: 设这个窃贼需要走Xh 才能到达地面，并设Y 代表他每次对3 扇门的选择情况，Y 各以$\frac{1}{3}$ 的概率取值1,2,3. 则

$$\mathrm{E}(X) = \mathrm{E}[E(X|Y)] = \sum_{i=1}^{3} \mathrm{E}(X|Y=i)\mathrm{P}(Y=i),$$

注意到$\mathrm{E}(X|Y=1)=3$, $\mathrm{E}(X|Y=2)=5+\mathrm{E}(X)$, $\mathrm{E}(X|Y=3)=7+\mathrm{E}(X)$, 所以

$$\mathrm{E}(X) = \frac{1}{3}[3+5+\mathrm{E}(X)+7+E(X)].$$

即得到$\mathrm{E}(X)=15$.

§3.6 随机变量的方差

§3.6 随机变量的方差

在许多问题中，只知道随机变量的数学期望是远远不够的. 在例1.2中，如果知道全班期中考试的平均成绩是75分，我们还是不知道这个班的学习情况是否整齐，也不知道这次考试的试题是否合理.

可以设想，当全班的成绩过于集中时，试题可能有问题. 过于分散时，会有较多的学生不及格，也不合乎情理. 于是应当有一个量描述考试成绩是否过于集中或分散. 这个量就是方差.

§3.6 随机变量的方差

§3.6 随机变量的方差

在许多问题中，只知道随机变量的数学期望是远远不够的. 在例1.2中，如果知道全班期中考试的平均成绩是75分，我们还是不知道这个班的学习情况是否整齐，也不知道这次考试的试题是否合理.

可以设想，当全班的成绩过于集中时，试题可能有问题. 过于分散时，会有较多的学生不及格，也不合乎情理. 于是应当有一个量描述考试成绩是否过于集中或分散. 这个量就是方差.

§3.6 随机变量的方差

例3.6.1 在例1.2中，全班同学的期中考试的平均分是

$$\mu = \frac{1}{m}\sum_{i=1}^{m} x_i,\ \ m = 180.$$

可以用

$$\sigma^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)^2$$

描述全班期中考试成绩的分散程度．因为考试成绩是总体，所以称σ^2是总体方差．总体方差用来描述总体的分散程度．

§3.6 随机变量的方差

例3.6.1 在例1.2中，全班同学的期中考试的平均分是

$$\mu = \frac{1}{m}\sum_{i=1}^{m} x_i,\ \ m = 180.$$

可以用

$$\sigma^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)^2$$

描述全班期中考试成绩的分散程度. 因为考试成绩是总体，所以称σ^2是总体方差. 总体方差用来描述总体的分散程度.

§3.6 随机变量的方差

从例3.1.2知道，若用X表示任选一个同学的期中成绩，则X的概率分布(1.1)是总体分布， $\mu = \mathrm{E}\,X$是总体均值. 因为X的取值的分散程度由总体的分散程度σ^2决定，所以把X的方差定义成总体方差σ^2. 因为得j分的同学共有m_j个， $p_j = P(X = j) = m_j/m$, 所以有

$$\begin{aligned}\sigma^2 &= \frac{1}{m}\sum_{i=1}^{m}(x_i-\mu)^2\\ &= \sum_{j=0}^{100}(j-\mu)^2\frac{m_j}{m}\\ &= \sum_{j=0}^{100}(j-\mu)^2 p_j\\ &= \mathrm{E}\,(X-\mu)^2.\end{aligned}$$

说明应当用$\mathrm{E}\,(X-\mathrm{E}\,X)^2$ 描述随机变量X的分散程度.

§3.6 随机变量的方差

从例3.1.2知道，若用X表示任选一个同学的期中成绩，则X的概率分布(1.1)是总体分布， $\mu = \mathrm{E}\,X$是总体均值. 因为X的取值的分散程度由总体的分散程度σ^2决定，所以把X的方差定义成总体方差σ^2. 因为得j分的同学共有m_j个， $p_j = P(X = j) = m_j/m$, 所以有

$$\begin{aligned}\sigma^2 &= \frac{1}{m}\sum_{i=1}^{m}(x_i-\mu)^2\\ &= \sum_{j=0}^{100}(j-\mu)^2\frac{m_j}{m}\\ &= \sum_{j=0}^{100}(j-\mu)^2 p_j\\ &= \mathrm{E}\,(X-\mu)^2.\end{aligned}$$

说明应当用$\mathrm{E}\,(X-\mathrm{E}\,X)^2$ 描述随机变量X的分散程度.

§3.5 随机变量的方差

定义3.6.1 设$\mu = \mathrm{E}\,X$, 如果$\mathrm{E}\,(X-\mu)^2 < \infty$, 则称

$$\sigma^2 = \mathrm{E}\,(X-\mu)^2 \qquad (6.1)$$

为X的方差(variance), 记做$\mathrm{Var}\,(X)$ 或σ_{XX}.
称$\sigma_X = \sqrt{\mathrm{Var}\,(X)}$为$X$的标准差.

也可以从另外的角度解释方差. 设X是对长度为μ的物体的测量值, 则$X-\mu$是测量误差, $(X-\mu)^2$是测量误差的平方. 如果测量仪器无系统偏差(即$\mathrm{E}\,X=\mu$), 则$\mathrm{E}\,(X-\mu)^2$是测量误差平方的平均, 正是方差.

§3.5 随机变量的方差

定义3.6.1 设$\mu = \mathrm{E}\,X$, 如果$\mathrm{E}\,(X-\mu)^2 < \infty$, 则称

$$\sigma^2 = \mathrm{E}\,(X-\mu)^2 \qquad (6.1)$$

为X的方差(variance), 记做$\mathrm{Var}\,(X)$ 或σ_{XX}.
称$\sigma_X = \sqrt{\mathrm{Var}\,(X)}$为$X$的标准差.

也可以从另外的角度解释方差. 设X是对长度为μ的物体的测量值, 则$X-\mu$是测量误差, $(X-\mu)^2$是测量误差的平方. 如果测量仪器无系统偏差(即$\mathrm{E}\,X = \mu$), 则$\mathrm{E}\,(X-\mu)^2$是测量误差平方的平均, 正是方差.

§3.6 随机变量的方差

当X有离散分布$p_j = P(X = x_j), j = 1, 2, \cdots$时, 利用定理3.3.2(1)得到

$$\mathrm{Var}\,(X) = \mathrm{E}\,(X - \mu)^2 = \sum_{j=1}^{\infty}(x_j - \mu)^2 p_j.$$

当X有概率密度$f(x)$时, 利用定理3.3.1(1)得到

$$\mathrm{Var}\,(X) = \mathrm{E}\,(X - \mu)^2 = \int_{-\infty}^{\infty}(x - \mu)^2 f(x)\,\mathrm{d}x.$$

上面两式都说明随机变量X的方差$\mathrm{Var}\,(X)$由X的概率分布唯一决定. 这就是下面的定理.

§3.6　随机变量的方差

当X有离散分布$p_j = P(X = x_j), j = 1, 2, \cdots$时，利用定理3.3.2(1)得到

$$\mathrm{Var}\,(X) = \mathrm{E}\,(X - \mu)^2 = \sum_{j=1}^{\infty} (x_j - \mu)^2 p_j.$$

当X有概率密度$f(x)$时，利用定理3.3.1(1)得到

$$\mathrm{Var}\,(X) = \mathrm{E}\,(X - \mu)^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\,\mathrm{d}x.$$

上面两式都说明随机变量X的方差$\mathrm{Var}\,(X)$由X的概率分布唯一决定．这就是下面的定理．

§3.6 随机变量的方差

定理3.6.1 如果X, Y有相同的概率分布, 则他们有相同的数学期望和方差.

X的方差描述了X的分散程度, $\mathrm{Var}(X)$越小, 说明X在数学期望μ附近越集中. 特别当$\mathrm{Var}(X)=0$时, 知道$X=\mu$ a.s..
利用方差的定义得到

$$\mathrm{Var}(X)=\mathrm{E}(X-\mu)^2=\mathrm{E}(X^2-2X\mu+\mu^2)=\mathrm{E}X^2-\mu^2.$$

这就得到计算方差的常用公式

$$\mathrm{Var}(X)=\mathrm{E}X^2-(\mathrm{E}X)^2. \qquad (3.6.2)$$

§3.6 随机变量的方差

定理3.6.1 如果X, Y有相同的概率分布, 则他们有相同的数学期望和方差.

X的方差描述了X的分散程度, $\mathrm{Var}\,(X)$越小, 说明X在数学期望μ附近越集中. 特别当$\mathrm{Var}\,(X)=0$时, 知道$X=\mu$ a.s..
利用方差的定义得到

$$\mathrm{Var}\,(X)=\mathrm{E}\,(X-\mu)^2=\mathrm{E}\,(X^2-2X\mu+\mu^2)=\mathrm{E}\,X^2-\mu^2.$$

这就得到计算方差的常用公式

$$\mathrm{Var}\,(X)=\mathrm{E}\,X^2-(\mathrm{E}\,X)^2. \qquad (3.6.2)$$

§3.6 随机变量的方差

A. 常用的方差

下面计算几个常见分布的方差.

(1) 伯努利分布$\mathcal{B}(1, p)$:

设$P(X=1)=p$, $P(X=0)=1-p=q$, 则$\mathrm{Var}\,(X)=pq$.

证明 由$X^2=X$ 和$\mathrm{E}\,X=p$, 得到

$$\mathrm{Var}\,(X)=\mathrm{E}\,X^2-(\mathrm{E}\,X)^2=p-p^2=pq.$$

§3.6 随机变量的方差

A. 常用的方差

下面计算几个常见分布的方差.

(1) 伯努利分布$\mathcal{B}(1,p)$:

设$P(X=1)=p$, $P(X=0)=1-p=q$, 则$\mathrm{Var}\,(X)=pq$.

证明 由$X^2=X$ 和$\mathrm{E}\,X=p$, 得到

$$\mathrm{Var}\,(X)=\mathrm{E}\,X^2-(\mathrm{E}\,X)^2=p-p^2=pq.$$

§3.6 随机变量的方差

(2) 二项分布$\mathcal{B}(n,p)$: 设$q=1-p$,

$$P(X=j)=\mathrm{C}_n^j p^j q^{n-j},\ 0\leqslant j\leqslant n,$$

则$\mathrm{Var}\,(X)=npq$.

证明 由$\mathrm{E}\,X=np$ 和$\mathrm{E}\,[X(X-1)]=n(n-1)p^2$ 得到

$$\mathrm{E}\,X^2=\ \mathrm{E}[X(X-1)]+\mathrm{E}\,X=\ n(n-1)p^2+np.$$

则

$$\mathrm{Var}\,(X)=n(n-1)p^2+np-(np)^2=npq.$$

§3.6 随机变量的方差

(2) 二项分布$\mathcal{B}(n,p)$： 设$q=1-p$,

$$P(X=j)=\mathrm{C}_n^j p^j q^{n-j},\ 0\leqslant j\leqslant n,$$

则$\mathrm{Var}\,(X)=npq$.

证明 由$\mathrm{E}\,X=np$ 和$\mathrm{E}\,[X(X-1)]=n(n-1)p^2$ 得到

$$\mathrm{E}\,X^2=\mathrm{E}[X(X-1)]+\mathrm{E}\,X=n(n-1)p^2+np.$$

则

$$\mathrm{Var}\,(X)=n(n-1)p^2+np-(np)^2=npq.$$

§3.6 随机变量的方差

(2) 二项分布$\mathcal{B}(n,p)$: 设$q=1-p$,

$$P(X=j)=\mathrm{C}_n^j p^j q^{n-j},\ 0\leqslant j\leqslant n,$$

则$\mathrm{Var}\,(X)=npq$.

证明 由$\mathrm{E}\,X=np$ 和$\mathrm{E}\,[X(X-1)]=n(n-1)p^2$ 得到

$$\mathrm{E}\,X^2=\ \mathrm{E}[X(X-1)]+\mathrm{E}\,X=\ n(n-1)p^2+np.$$

则

$$\mathrm{Var}\,(X)=n(n-1)p^2+np-(np)^2=npq.$$

§3.6 随机变量的方差

(3) 泊松分布$\mathcal{P}(\lambda)$: 设

$$P(X=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,\cdots,$$

则$\mathrm{Var}\,(X)=\lambda.$

证明 由$\mathrm{E}\,X=\lambda$ 得到

$$\begin{aligned}\mathrm{E}\,X^2 &= \mathrm{E}\,[X(X-1)]+\mathrm{E}\,X\\ &= \sum_{k=0}^{\infty}k(k-1)\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}+\lambda\\ &= \lambda^2\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}\mathrm{e}^{-\lambda}+\lambda=\lambda^2+\lambda.\end{aligned}$$

则

$$\mathrm{Var}\,(X)=\lambda^2+\lambda-\lambda^2=\lambda.$$

§3.6 随机变量的方差

(3) 泊松分布$\mathcal{P}(\lambda)$: 设

$$P(X=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,\cdots,$$

则$\mathrm{Var}\,(X)=\lambda$.

证明 由$\mathrm{E}\,X=\lambda$ 得到

$$\begin{aligned}\mathrm{E}\,X^2 &= \mathrm{E}\,[X(X-1)]+\mathrm{E}\,X\\ &= \sum_{k=0}^{\infty}k(k-1)\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}+\lambda\\ &= \lambda^2\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}\mathrm{e}^{-\lambda}+\lambda=\lambda^2+\lambda.\end{aligned}$$

则

$$\mathrm{Var}\,(X)=\lambda^2+\lambda-\lambda^2=\lambda.$$

§3.6 随机变量的方差

(3) 泊松分布$\mathcal{P}(\lambda)$: 设

$$P(X=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,\cdots,$$

则$\mathrm{Var}\,(X)=\lambda.$

证明 由$\mathrm{E}\,X=\lambda$ 得到

$$\begin{aligned}\mathrm{E}\,X^2 &= \mathrm{E}\,[X(X-1)]+\mathrm{E}\,X\\ &= \sum_{k=0}^{\infty}k(k-1)\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}+\lambda\\ &= \lambda^2\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}\mathrm{e}^{-\lambda}+\lambda=\lambda^2+\lambda.\end{aligned}$$

则

$$\mathrm{Var}\,(X)=\lambda^2+\lambda-\lambda^2=\lambda.$$

§3.6 随机变量的方差

(4) 几何分布: 设X有概率分布

$$P(X=j)=pq^{j-1},\ j=1,2,\cdots,\ q=1-p,$$

则$\mathrm{Var}\,(X)=q/p^2$.

证明 用$\mathrm{E}\,X=1/p$得到

§3.6 随机变量的方差

(4) 几何分布: 设X有概率分布

$$P(X=j)=pq^{j-1},\ j=1,2,\cdots,\ q=1-p,$$

则$\mathrm{Var}\,(X)=q/p^2$.

证明 用$\mathrm{E}\,X=1/p$得到

§3.6 随机变量的方差

(4) 几何分布: 设X有概率分布

$$P(X=j)=pq^{j-1},\ j=1,2,\cdots,\ q=1-p,$$

则$\mathrm{Var}\,(X)=q/p^2$.

证明　用$\mathrm{E}\,X=1/p$得到

§3.6 随机变量的方差

$$
\begin{aligned}
\mathrm{E}\,X^2 &= \mathrm{E}\,[X(X-1)] + \mathrm{E}\,X \\
&= \sum_{j=1}^{\infty} j(j-1)pq^{j-1} + \frac{1}{p} \\
&= pq\Big(\sum_{j=0}^{\infty} q^j\Big)'' + \frac{1}{p} \\
&= pq\Big(\frac{1}{1-q}\Big)'' + \frac{1}{p} \\
&= \frac{2pq}{(1-q)^3} + \frac{1}{p} \\
&= \frac{2q}{p^2} + \frac{1}{p}.
\end{aligned}
$$

$$
\mathrm{Var}\,(X) = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{q}{p^2}.
$$

§3.6 随机变量的方差

(5) 均匀分布$U(a,b)$: 设X有概率密度 $f(x)=1/(b-a)$, $x\in(a,b)$, 则

$$\mathrm{Var}\,(X)=\frac{(b-a)^2}{12}.$$

证明 因为X有数学期望$\mathrm{E}\,X=(a+b)/2$, 且

$$\mathrm{E}\,X^2=\int_a^b\frac{x^2}{b-a}\,\mathrm{d}x=\frac{b^3-a^3}{3(b-a)}.$$

所以

$$\mathrm{Var}\,(X)=\frac{b^3-a^3}{3(b-a)}-\Big(\frac{a+b}{2}\Big)^2=\frac{(b-a)^2}{12}.$$

§3.6 随机变量的方差

(5) 均匀分布$\mathcal{U}(a,b)$: 设X有概率密度 $f(x)=1/(b-a)$, $x\in(a,b)$, 则

$$\mathrm{Var}\,(X)=\frac{(b-a)^2}{12}.$$

证明 因为X有数学期望$\mathrm{E}\,X=(a+b)/2$, 且

$$\mathrm{E}\,X^2=\int_a^b\frac{x^2}{b-a}\,\mathrm{d}x=\frac{b^3-a^3}{3(b-a)}.$$

所以

$$\mathrm{Var}\,(X)=\frac{b^3-a^3}{3(b-a)}-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}.$$

§3.6 随机变量的方差

(5) 均匀分布$U(a,b)$: 设X有概率密度 $f(x)=1/(b-a)$, $x\in(a,b)$, 则

$$\operatorname{Var}(X)=\frac{(b-a)^2}{12}.$$

证明 因为X有数学期望$\mathrm{E}\,X=(a+b)/2$, 且

$$\mathrm{E}\,X^2=\int_a^b\frac{x^2}{b-a}\,\mathrm{d}x=\frac{b^3-a^3}{3(b-a)}.$$

所以

$$\operatorname{Var}(X)=\frac{b^3-a^3}{3(b-a)}-\Big(\frac{a+b}{2}\Big)^2=\frac{(b-a)^2}{12}.$$

§3.6 随机变量的方差

(6) 指数分布$\mathcal{E}(\lambda)$: 设X有概率密度$f(x)=\lambda e^{-\lambda x}$, $x>0$, 则$\mathrm{Var}\,(X)=1/\lambda^2$.

证明 X有数学期望$\mathrm{E}\,X=1/\lambda$, 由

$$\begin{aligned}
\mathrm{E}\,X^2 &= \int_0^{\infty} x^2\lambda e^{-\lambda x}\,\mathrm{d}x \\
&= \frac{1}{\lambda^2}\int_0^{\infty} t^2 e^{-t}\,\mathrm{d}t \qquad [\text{取}x=t/\lambda] \\
&= \frac{1}{\lambda^2}\Gamma(3) \\
&= \frac{2!}{\lambda^2},
\end{aligned}$$

得到

$$\mathrm{Var}\,(X)=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

§3.6 随机变量的方差

(6) 指数分布$\mathcal{E}(\lambda)$: 设X有概率密度$f(x)=\lambda \mathrm{e}^{-\lambda x}$, $x>0$, 则$\mathrm{Var}\,(X)=1/\lambda^2$.

证明 X有数学期望$\mathrm{E}\,X=1/\lambda$, 由

$$\begin{aligned}
\mathrm{E}\,X^2 &= \int_0^\infty x^2\lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x \\
&= \frac{1}{\lambda^2}\int_0^\infty t^2\mathrm{e}^{-t}\,\mathrm{d}t \qquad [\text{取}x=t/\lambda] \\
&= \frac{1}{\lambda^2}\Gamma(3) \\
&= \frac{2!}{\lambda^2},
\end{aligned}$$

得到

$$\mathrm{Var}\,(X)=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

§3.6 随机变量的方差

(6) 指数分布$\mathcal{E}(\lambda)$: 设X有概率密度$f(x)=\lambda\mathrm{e}^{-\lambda x}$, $x>0$, 则$\mathrm{Var}\,(X)=1/\lambda^2$.

证明 X有数学期望$\mathrm{E}\,X=1/\lambda$, 由

$$\begin{aligned}\mathrm{E}\,X^2 &= \int_0^\infty x^2\lambda\mathrm{e}^{-\lambda x}\,\mathrm{d}x\\ &= \frac{1}{\lambda^2}\int_0^\infty t^2\mathrm{e}^{-t}\,\mathrm{d}t \qquad [\text{取}x=t/\lambda]\\ &= \frac{1}{\lambda^2}\Gamma(3)\\ &= \frac{2!}{\lambda^2},\end{aligned}$$

得到

$$\mathrm{Var}\,(X)=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

§3.6 随机变量的方差

(7) 正态分布$N(\mu,\sigma^2)$: 设$X \sim N(\mu,\sigma^2)$, 则$\mathrm{Var}\,(X)=\sigma^2$.

证明 X有数学期望$\mathrm{E}\,X=\mu$, 并且由§3.4例3.4.2知道

$$Y=\frac{X-\mu}{\sigma}\sim N(0,1).$$

从例3.4.1知道$\mathrm{E}\,Y^2=1$, 于是用$(X-\mu)^2=Y^2\sigma^2$ 得到

$$\mathrm{Var}\,(X)=\mathrm{E}\,(X-\mu)^2=\mathrm{E}\,Y^2\sigma^2=\sigma^2.$$

现在我们知道了正态分布$N(\mu,\sigma^2)$中的μ和σ^2就是该正态分布的数学期望和方差. 如果已知X服从正态分布, 那么只要再计算它的数学期望μ和方差σ^2, 则可以得到X 的概率密度了.

§3.6 随机变量的方差

(7) 正态分布$N(\mu,\sigma^2)$: 设$X\sim N(\mu,\sigma^2)$, 则$\mathrm{Var}\,(X)=\sigma^2$.

证明 X有数学期望$\mathrm{E}\,X=\mu$, 并且由§3.4例3.4.2知道

$$Y=\frac{X-\mu}{\sigma}\sim N(0,1).$$

从例3.4.1知道$\mathrm{E}\,Y^2=1$, 于是用$(X-\mu)^2=Y^2\sigma^2$ 得到

$$\mathrm{Var}\,(X)=\mathrm{E}\,(X-\mu)^2=\mathrm{E}\,Y^2\sigma^2=\sigma^2.$$

现在我们知道了正态分布$N(\mu,\sigma^2)$中的μ和σ^2就是该正态分布的数学期望和方差. 如果已知X服从正态分布, 那么只要再计算它的数学期望μ和方差σ^2, 则可以得到X 的概率密度了.

§3.6 随机变量的方差

例3.6.2 设X, Y相互独立, 都服从标准正态分布, 求

$$U = 3X - 4Y + 5$$

的分布.

解 U服从正态分布. 再从$\mathrm{E}\,X = \mathrm{E}\,Y = 0$, $\mathrm{Var}\,(X) = \mathrm{Var}\,(Y) = 1$ 得到$\mathrm{E}\,X^2 = \mathrm{E}\,Y^2 = 1$, 从而得到

$$\begin{aligned}
\mathrm{E}\,U &= 3\mathrm{E}\,X - 4\mathrm{E}\,Y + 5 = 5, \\
\mathrm{E}\,(XY) &= \mathrm{E}\,X \cdot \mathrm{E}\,Y = 0, \\
\mathrm{Var}\,(U) &= \mathrm{E}\,(U - \mathrm{E}\,U)^2 = \mathrm{E}\,(3X - 4Y)^2 \\
&= 9\mathrm{E}\,X^2 + 16\mathrm{E}\,Y^2 - 24\mathrm{E}\,(XY) \\
&= 25.
\end{aligned}$$

于是$U \sim N(5, 25)$.

§3.6 随机变量的方差

例3.6.2 设X, Y相互独立, 都服从标准正态分布, 求

$$U = 3X - 4Y + 5$$

的分布.

解 U服从正态分布. 再从$\mathrm{E}\,X = \mathrm{E}\,Y = 0$, $\mathrm{Var}\,(X) = \mathrm{Var}\,(Y) = 1$ 得到$\mathrm{E}\,X^2 = \mathrm{E}\,Y^2 = 1$, 从而得到

$$\begin{aligned}
\mathrm{E}\,U &= 3\mathrm{E}\,X - 4\mathrm{E}\,Y + 5 = 5,\\
\mathrm{E}\,(XY) &= \mathrm{E}\,X \cdot \mathrm{E}\,Y = 0,\\
\mathrm{Var}\,(U) &= \mathrm{E}\,(U - \mathrm{E}\,U)^2 = \mathrm{E}\,(3X - 4Y)^2\\
&= 9\mathrm{E}\,X^2 + 16\mathrm{E}\,Y^2 - 24\mathrm{E}\,(XY)\\
&= 25.
\end{aligned}$$

于是$U \sim N(5, 25)$.

§3.6 随机变量的方差

例3.6.2 设X, Y相互独立，都服从标准正态分布，求

$$U = 3X - 4Y + 5$$

的分布.

解 U服从正态分布. 再从$\mathrm{E}\,X = \mathrm{E}\,Y = 0$, $\mathrm{Var}\,(X) = \mathrm{Var}\,(Y) = 1$ 得到$\mathrm{E}\,X^2 = \mathrm{E}\,Y^2 = 1$, 从而得到

$$\begin{aligned} \mathrm{E}\,U &= 3\mathrm{E}\,X - 4\mathrm{E}\,Y + 5 = 5, \\ \mathrm{E}\,(XY) &= \mathrm{E}\,X \cdot \mathrm{E}\,Y = 0, \\ \mathrm{Var}\,(U) &= \mathrm{E}\,(U - \mathrm{E}\,U)^2 = \mathrm{E}\,(3X - 4Y)^2 \\ &= 9\mathrm{E}\,X^2 + 16\mathrm{E}\,Y^2 - 24\mathrm{E}\,(XY) \\ &= 25. \end{aligned}$$

于是$U \sim N(5, 25)$.

§3.6 随机变量的方差

B. 方差的性质

定理3.6.2 设a, b, c 是常数, $\mathrm{E}\,X = \mu$, $\mathrm{Var}\,(X) < \infty$, $\mu_j = \mathrm{E}\,X_j$, $\mathrm{Var}\,(X_j) < \infty (1 \leqslant j \leqslant n)$, 则

(1) $\mathrm{Var}\,(a + bX) = b^2 \mathrm{Var}\,(X)$;

(2) $\mathrm{Var}\,(X) = \mathrm{E}\,(X - \mu)^2 < \mathrm{E}\,(X - c)^2$, 只要常数$c \neq \mu$;

(3) $\mathrm{Var}\,(X) = 0$的充分必要条件是$P(X = \mu) = 1$;

(4) 当$X_1, X_2, \cdots, X_n$相互独立时,

$$\mathrm{Var}\,\Big(\sum_{j=1}^{n} X_j\Big) = \sum_{j=1}^{n} \mathrm{Var}\,(X_j).$$

§3.6 随机变量的方差

B. 方差的性质

定理3.6.2 设a, b, c 是常数, $\mathrm{E}\,X=\mu$, $\mathrm{Var}\,(X)<\infty$, $\mu_j=\mathrm{E}\,X_j$, $\mathrm{Var}\,(X_j)<\infty(1\leqslant j\leqslant n)$, 则

(1) $\mathrm{Var}\,(a+bX)=b^2\mathrm{Var}\,(X)$;

(2) $\mathrm{Var}\,(X)=\mathrm{E}\,(X-\mu)^2<\mathrm{E}\,(X-c)^2$, 只要常数$c\neq\mu$;

(3) $\mathrm{Var}\,(X)=0$的充分必要条件是$P(X=\mu)=1$;

(4) 当$X_1, X_2, \cdots, X_n$相互独立时,

$$\mathrm{Var}\,\Big(\sum_{j=1}^{n}X_j\Big)=\sum_{j=1}^{n}\mathrm{Var}\,(X_j).$$

§3.6 随机变量的方差

*证明 (1) 由方差的定义得到

$$\begin{aligned}\mathrm{Var}\,(a+bX) &= \mathrm{E}\left[a+bX-(a+b\mathrm{E}\,X)\right]^2\\ &= \mathrm{E}\,[b^2(X-\mu)^2]\\ &= b^2\mathrm{Var}\,(X).\end{aligned}$$

(2) 对$c\neq\mu$, 由$\mathrm{E}\,(X-\mu)=0$得到

$$\begin{aligned}\mathrm{E}\,(X-c)^2 &= \mathrm{E}\,(X-\mu+\mu-c)^2\\ &= \mathrm{E}\,(X-\mu)^2+2\mathrm{E}\,(X-\mu)(\mu-c)+\mathrm{E}\,(\mu-c)^2\\ &= \mathrm{Var}\,(X)+(\mu-c)^2>\mathrm{Var}\,(X).\end{aligned}$$

于是结论(2)成立.

§3.6 随机变量的方差

*证明　(1) 由方差的定义得到

$$
\begin{aligned}
\mathrm{Var}\,(a+bX) &= \mathrm{E}\left[a+bX-(a+b\mathrm{E}\,X)\right]^2 \\
&= \mathrm{E}\,[b^2(X-\mu)^2] \\
&= b^2\mathrm{Var}\,(X).
\end{aligned}
$$

(2) 对$c \neq \mu$, 由$\mathrm{E}\,(X-\mu)=0$得到

$$
\begin{aligned}
\mathrm{E}\,(X-c)^2 &= \mathrm{E}\,(X-\mu+\mu-c)^2 \\
&= \mathrm{E}\,(X-\mu)^2+2\mathrm{E}\,(X-\mu)(\mu-c)+\mathrm{E}\,(\mu-c)^2 \\
&= \mathrm{Var}\,(X)+(\mu-c)^2 > \mathrm{Var}\,(X).
\end{aligned}
$$

于是结论(2)成立.

§3.6 随机变量的方差

*证明 (1) 由方差的定义得到

$$\begin{aligned}\mathrm{Var}\,(a+bX) &= \mathrm{E}\left[a+bX-(a+b\mathrm{E}\,X)\right]^2\\ &= \mathrm{E}\,[b^2(X-\mu)^2]\\ &= b^2\mathrm{Var}\,(X).\end{aligned}$$

(2) 对$c\neq\mu$, 由$\mathrm{E}\,(X-\mu)=0$得到

$$\begin{aligned}\mathrm{E}\,(X-c)^2 &= \mathrm{E}\,(X-\mu+\mu-c)^2\\ &= \mathrm{E}\,(X-\mu)^2+2\mathrm{E}\,(X-\mu)(\mu-c)+\mathrm{E}\,(\mu-c)^2\\ &= \mathrm{Var}\,(X)+(\mu-c)^2>\mathrm{Var}\,(X).\end{aligned}$$

于是结论(2)成立.

§3.6 随机变量的方差

(3) 如果$\mathrm{E}\,(X-\mu)^2=\mathrm{Var}\,(X)=0$，则由定理4.2的结论得到$(X-\mu)^2=0$ a.s.，即$X=\mu$ a.s.. 如果$P(X=\mu)=1$，则按定义$\mathrm{E}\,(X-\mu)^2=1\cdot(\mu-\mu)^2=0$.

(4)的证明留给读者.

在性质(1)中取$b=1$得到$\mathrm{Var}\,(a+X)=\mathrm{Var}\,(X)$，说明对随机变量进行常数平移后，随机变量的分散程度不变；取$a=0$得到$\mathrm{Var}\,(bX)=b^2\mathrm{Var}\,(X)$，说明将$X$扩大$b$倍后，标准差扩大$|b|$倍. (2)说明随机变量$X$在均方误差的意义下距离数学期望$\mu$最近. (3)说明除了以概率1等于常数的随机变量外，任何随机变量的方差都大于零.

以后无特殊说明时，都认为所述随机变量的方差大于零.

§3.6 随机变量的方差

(3) 如果$\mathrm{E}(X-\mu)^2=\mathrm{Var}(X)=0$, 则由定理4.2的结论得到$(X-\mu)^2=0$ a.s., 即$X=\mu$ a.s.. 如果$P(X=\mu)=1$, 则按定义$\mathrm{E}(X-\mu)^2=1\cdot(\mu-\mu)^2=0$.

(4)的证明留给读者.

在性质(1)中取$b=1$得到$\mathrm{Var}(a+X)=\mathrm{Var}(X)$, 说明对随机变量进行常数平移后, 随机变量的分散程度不变; 取$a=0$得到$\mathrm{Var}(bX)=b^2\mathrm{Var}(X)$, 说明将$X$扩大$b$倍后, 标准差扩大$|b|$倍. (2)说明随机变量$X$在均方误差的意义下距离数学期望$\mu$最近. (3)说明除了以概率1等于常数的随机变量外, 任何随机变量的方差都大于零.

以后无特殊说明时, 都认为所述随机变量的方差大于零.

§3.6 随机变量的方差

(3) 如果$\mathrm{E}(X-\mu)^2=\mathrm{Var}(X)=0$, 则由定理4.2的结论得到$(X-\mu)^2=0$ a.s., 即$X=\mu$ a.s.. 如果$P(X=\mu)=1$, 则按定义$\mathrm{E}(X-\mu)^2=1\cdot(\mu-\mu)^2=0$.

(4)的证明留给读者.

在性质(1)中取$b=1$得到$\mathrm{Var}(a+X)=\mathrm{Var}(X)$, 说明对随机变量进行常数平移后, 随机变量的分散程度不变; 取$a=0$得到$\mathrm{Var}(bX)=b^2\mathrm{Var}(X)$, 说明将$X$扩大$b$倍后, 标准差扩大$|b|$ 倍. (2)说明随机变量X 在均方误差的意义下距离数学期望μ最近. (3) 说明除了以概率1等于常数的随机变量外, 任何随机变量的方差都大于零.

以后无特殊说明时, 都认为所述随机变量的方差大于零.

§3.6 随机变量的方差

设$\sigma^2 = \mathrm{Var}\,(X) < \infty$，$Y = (X - \mathrm{E}\,X)/\sigma$，则

$$\mathrm{E}\,Y = 0,\ \ \mathrm{Var}\,(Y) = \frac{1}{\sigma^2}\mathrm{Var}\,(X - \mathrm{E}\,X) = 1.$$

这时称Y是X的标准化. 特别地，当$X \sim N(\mu, \sigma^2)$，$Y \sim N(0, 1)$.

例3.6.3 设$X_1, X_2, \cdots, X_n$相互独立，有共同的方差$\sigma^2 < \infty$，则

$$\mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) = \frac{1}{n}\sigma^2.$$

§3.6 随机变量的方差

设$\sigma^2 = \mathrm{Var}\,(X) < \infty$, $Y = (X - \mathrm{E}\,X)/\sigma$, 则

$$\mathrm{E}\,Y = 0,\ \mathrm{Var}\,(Y) = \frac{1}{\sigma^2}\mathrm{Var}\,(X - \mathrm{E}\,X) = 1.$$

这时称Y是X的标准化. 特别地, 当$X \sim N(\mu, \sigma^2)$, $Y \sim N(0, 1)$.

例3.6.3 设$X_1, X_2, \cdots, X_n$相互独立, 有共同的方差$\sigma^2 < \infty$, 则

$$\mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) = \frac{1}{n}\sigma^2.$$

§3.6 随机变量的方差

证明　用定理3.6.2的(1)和(4)得到

$$\begin{aligned}
\mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) &= \frac{1}{n^2}\mathrm{Var}\left(\sum_{j=1}^{n} X_j\right) \\
&= \frac{1}{n^2}\sum_{j=1}^{n}\mathrm{Var}(X_j) \\
&= \frac{1}{n^2}\sum_{j=1}^{n}\sigma^2 = \frac{1}{n}\sigma^2.
\end{aligned}$$

§3.6 随机变量的方差

在例3.6.3中，如果X_i是第i次测量重量为μ的物体时的测量值，测量的均方误差是$\mathrm{Var}(X_i)=\sigma^2$. 当用$n$次测量的平均

$$\overline{X}_n=\frac{1}{n}\sum_{j=1}^{n}X_j$$

作为μ的测量值时，方差降低n倍. 说明只要测量仪器没有系统偏差(指$\mathrm{E}\,X=\mu$，这时有$\mathrm{E}\,\overline{X}_n=\mu$)，测量精度总可以通过多次测量的平均来改进.

下面是用方差的性质计算方差的例子.

§3.6 随机变量的方差

例3.6.4　设$X \sim \mathcal{B}(n,p)$，则$\mathrm{Var}\,(X) = npq$.

证明　设$X_1, X_2, \cdots, X_n$独立同分布，都服从伯努利分布$\mathcal{B}(1,p)$，则$Y = X_1 + X_2 + \cdots + X_n \sim \mathcal{B}(n,p)$. 利用$\mathrm{Var}\,(X_i) = pq$ 得到

$$\begin{aligned}\mathrm{Var}\,(X) &= \mathrm{Var}\,(Y)\\ &= \mathrm{Var}\,(X_1) + \mathrm{Var}\,(X_2) + \cdots + \mathrm{Var}\,(X_n)\\ &= npq.\end{aligned}$$

在金融领域，若用X表示某项投资的收益，则数学期望$\mathrm{E}\,X$是平均收益，而方差$\mathrm{Var}\,(X)$可以描述投资风险. 这是因为方差越大，收益X 的不确定性越大.

§3.6 随机变量的方差

例3.6.4 设$X \sim \mathcal{B}(n,p)$, 则$\mathrm{Var}\,(X) = npq$.

证明 设$X_1, X_2, \cdots, X_n$独立同分布, 都服从伯努利分布$\mathcal{B}(1,p)$, 则$Y = X_1 + X_2 + \cdots + X_n \sim \mathcal{B}(n,p)$. 利用$\mathrm{Var}\,(X_i) = pq$ 得到

$$\begin{aligned}\mathrm{Var}\,(X) &= \mathrm{Var}\,(Y)\\ &= \mathrm{Var}\,(X_1) + \mathrm{Var}\,(X_2) + \cdots + \mathrm{Var}\,(X_n)\\ &= npq.\end{aligned}$$

在金融领域, 若用X表示某项投资的收益, 则数学期望$\mathrm{E}\,X$是平均收益, 而方差$\mathrm{Var}\,(X)$可以描述投资风险. 这是因为方差越大, 收益X 的不确定性越大.

§3.7　协方差和相关系数

§3.7　协方差和相关系数

A. 内积不等式

定理3.7.1(内积不等式)　设$\mathrm{E}\,X^2 < \infty$, $\mathrm{E}\,Y^2 < \infty$, 则有

$$|\mathrm{E}\,(XY)| \leqslant \sqrt{\mathrm{E}\,X^2\,\mathrm{E}\,Y^2}, \qquad (7.1)$$

并且等号成立的充分必要条件是有不全为零的常数a, b, 使得

$$P(aX + bY = 0) = 1.$$

§3.7 协方差和相关系数

证明 对于不全为零的常数a, b, 二次型

$$\begin{aligned} \mathrm{E}\,(aX+bY)^2 &= a^2\mathrm{E}\,X^2 + 2ab\mathrm{E}\,(XY) + b^2\mathrm{E}\,Y^2 \\ &= (a,b)\boldsymbol{\Sigma}(a,b)^{\mathrm{T}} \geqslant 0, \qquad (6.2) \end{aligned}$$

其中

$$\boldsymbol{\Sigma} = \begin{pmatrix} \mathrm{E}\,X^2 & \mathrm{E}\,(XY) \\ \mathrm{E}\,(XY) & \mathrm{E}\,Y^2 \end{pmatrix}.$$

由$\boldsymbol{\Sigma}$的半正定性得到(7.1).
从$\det(\boldsymbol{\Sigma}) = \mathrm{E}\,X^2\,\mathrm{E}\,Y^2 - [\mathrm{E}\,(XY)]^2$, 知道(6.1)中的等号成立当且仅当$\boldsymbol{\Sigma}$退化, 当且仅当有不全为零的常数$a, b$使$\mathrm{E}\,(aX+bY)^2 = 0$, 当且仅当(见定理4.2)有不全为零的常数$a, b$使$P(aX+bY=0)=1$.

§3.7 协方差和相关系数

证明 对于不全为零的常数a,b, 二次型

$$\begin{aligned} \mathrm{E}\,(aX+bY)^2 = &\quad a^2\mathrm{E}\,X^2 + 2ab\mathrm{E}\,(XY) + b^2\mathrm{E}\,Y^2 \\ = &\quad (a,b)\boldsymbol{\Sigma}(a,b)^{\mathrm{T}} \geqslant 0, \qquad (6.2) \end{aligned}$$

其中

$$\boldsymbol{\Sigma} = \begin{pmatrix} \mathrm{E}\,X^2 & \mathrm{E}\,(XY) \\ \mathrm{E}\,(XY) & \mathrm{E}\,Y^2 \end{pmatrix}.$$

由$\boldsymbol{\Sigma}$的半正定性得到(7.1).
从$\det(\boldsymbol{\Sigma}) = \mathrm{E}\,X^2\,\mathrm{E}\,Y^2 - [\mathrm{E}\,(XY)]^2$, 知道(6.1)中的等号成立当且仅当$\boldsymbol{\Sigma}$退化, 当且仅当有不全为零的常数$a,b$使$\mathrm{E}\,(aX+bY)^2=0$, 当且仅当(见定理4.2)有不全为零的常数$a,b$使$P(aX+bY=0)=1$.

§3.7 协方差和相关系数

例3.7.1 若$\mathrm{E}\,X^2 < \infty$,则$\mathrm{E}\,|X| < \infty$.

证明 由内积不等式得到

$$\mathrm{E}\,|X| = \mathrm{E}\,|X| \cdot 1 \leqslant \sqrt{\mathrm{E}\,X^2 \mathrm{E}\,1^2} = \sqrt{\mathrm{E}\,X^2} < \infty.$$

§3.7 协方差和相关系数

B. 协方差和相关系数

为了研究随机变量X, Y的关系，引入协方差和相关系数的定义. 设$\sigma_X = \sqrt{\sigma_{XX}}$, $\sigma_Y = \sqrt{\sigma_{YY}}$分别是$X$, Y的标准差.

定义3.7.1 设$\mu_X = \mathrm{E}\,X$, $\mu_Y = \mathrm{E}\,Y$.

(1) 当σ_X, σ_Y存在时，称

$$\mathrm{E}\,[(X-\mu_X)(Y-\mu_Y)] \qquad (7.3)$$

为随机变量X, Y的协方差(covariance),记做$\mathrm{Cov}\,(X,Y)$ 或σ_{XY}. 当$\mathrm{Cov}\,(X,Y)=0$时，称X, Y 不相关.

(2) 当$0<\sigma_X\sigma_Y<\infty$, 称

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \qquad (7.4)$$

为X, Y的相关系数(correlation coefficient). 相关系数ρ_{XY}也常用$\rho(X,Y)$表示.

§3.7 协方差和相关系数

B. 协方差和相关系数

为了研究随机变量X, Y的关系，引入协方差和相关系数的定义. 设$\sigma_X = \sqrt{\sigma_{XX}}$, $\sigma_Y = \sqrt{\sigma_{YY}}$分别是$X$, Y的标准差.

定义3.7.1 设$\mu_X = \mathrm{E}\,X$, $\mu_Y = \mathrm{E}\,Y$.

(1) 当σ_X, σ_Y 存在时，称

$$\mathrm{E}\,[(X-\mu_X)(Y-\mu_Y)] \qquad (7.3)$$

为随机变量X, Y的协方差(covariance),记做$\mathrm{Cov}\,(X,Y)$ 或σ_{XY}. 当$\mathrm{Cov}\,(X,Y)=0$时，称X, Y 不相关.

(2) 当$0<\sigma_X\sigma_Y<\infty$, 称

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \qquad (7.4)$$

为X, Y的相关系数(correlation coefficient). 相关系数ρ_{XY}也常用$\rho(X,Y)$表示.

§3.7 协方差和相关系数

B. 协方差和相关系数

为了研究随机变量X, Y的关系，引入协方差和相关系数的定义. 设$\sigma_X = \sqrt{\sigma_{XX}}$, $\sigma_Y = \sqrt{\sigma_{YY}}$分别是$X$, Y的标准差.

定义3.7.1 设$\mu_X = \mathrm{E}\,X$, $\mu_Y = \mathrm{E}\,Y$.

(1) 当σ_X, σ_Y 存在时，称

$$\mathrm{E}\,[(X-\mu_X)(Y-\mu_Y)] \tag{7.3}$$

为随机变量X，Y的协方差(covariance),记做$\mathrm{Cov}\,(X, Y)$ 或σ_{XY}. 当$\mathrm{Cov}\,(X, Y) = 0$时，称X，Y 不相关.

(2) 当$0 < \sigma_X\sigma_Y < \infty$，称

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \tag{7.4}$$

为X，Y的相关系数(correlation coefficient). 相关系数ρ_{XY}也常用$\rho(X, Y)$表示.

§3.7 协方差和相关系数

下面是计算协方差的常用公式:

$$\sigma_{XY} = \mathrm{E}(XY) - (\mathrm{E}\,X)(\mathrm{E}\,Y). \qquad (7.5)$$

(7.5)的证明如下.

$$\begin{aligned}
\sigma_{XY} &= \mathrm{E}\,[(X-\mu_X)(Y-\mu_Y)] \\
&= \mathrm{E}\,(XY - \mu_X Y - X\mu_Y + \mu_X\mu_Y) \\
&= \mathrm{E}\,(XY) - \mu_X\mu_Y - \mu_X\mu_Y + \mu_X\mu_Y \\
&= \mathrm{E}\,(XY) - (\mathrm{E}\,X)(\mathrm{E}\,Y).
\end{aligned}$$

从(7.4)和内积不等式马上得到相关系数的性质如下.

§3.7 协方差和相关系数

定理3.7.2 设ρ_{XY} 是X, Y的相关系数, 则有

(1) $|\rho_{XY}| \leqslant 1$;

(2) $|\rho_{XY}| = 1$的充分必要条件是有常数a, b使得

$$P(Y = a + bX) = 1;$$

(3) 如果X, Y独立, 则X, Y不相关.

从定理3.7.2 (2)看出, 当$|\rho_{XY}| = 1$成立时, X, Y 有线性关系. 这时称 X, Y **线性相关**.

§3.7 协方差和相关系数

例3.7.2 设(X,Y)在单位圆$D=\{(x,y)|\,x^2+y^2\leqslant 1\}$ 内均匀分布，则X，Y不相关，也不独立.

证明 (X,Y) 有联合密度$f(x,y)=\mathrm{I}[D]/\pi$，其中示性函数

$$\mathrm{I}[D]=\mathrm{I}[\,|x|\leqslant\sqrt{1-y^2}\,]\cdot\mathrm{I}[\,|y|\leqslant 1].$$

注意x在$(-\sqrt{1-y^2},\sqrt{1-y^2}\,)$中的积分是0，用公式(3.2)得到

$$\begin{aligned}
\mathrm{E}\,X &= \iint_{\mathrm{R}^2} xf(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty}\mathrm{I}[D]x\,\mathrm{d}x\Big)\,\mathrm{d}y \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\Big(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x\,\mathrm{d}x\Big)\mathrm{I}[\,|y|\leqslant 1]\,\mathrm{d}y \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} 0\,\mathrm{I}[\,|y|\leqslant 1]\,\mathrm{d}y = 0.
\end{aligned}$$

§3.7 协方差和相关系数

例3.7.2 设(X,Y)在单位圆$D=\{(x,y)\,|\,x^2+y^2\leqslant 1\}$ 内均匀分布，则X，Y不相关，也不独立.

证明 (X,Y) 有联合密度$f(x,y)=\mathrm{I}[D]/\pi$，其中示性函数

$$\mathrm{I}[D]=\mathrm{I}[\,|x|\leqslant\sqrt{1-y^2}]\cdot\mathrm{I}[\,|y|\leqslant 1].$$

注意x在$(-\sqrt{1-y^2},\sqrt{1-y^2})$中的积分是0，用公式(3.2)得到

$$\begin{aligned}
\mathrm{E}X &= \iint_{\mathrm{R}^2} xf(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty}\mathrm{I}[D]x\,\mathrm{d}x\Big)\,\mathrm{d}y \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\Big(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}x\,\mathrm{d}x\Big)\mathrm{I}[\,|y|\leqslant 1]\,\mathrm{d}y \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}0\,\mathrm{I}[\,|y|\leqslant 1]\,\mathrm{d}y = 0.
\end{aligned}$$

§3.6 协方差和相关系数

对称地得到$\mathrm{E}\,Y=0$. 于是

$$\begin{aligned}\mathrm{Cov}\,(X,Y) &= \mathrm{E}\,(XY)\\ &= \iint_{\mathrm{R}^2} xyf(x,y)\,\mathrm{d}x\mathrm{d}y\\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} y\Big(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x\,\mathrm{d}x\Big)\mathrm{I}[\,|y|\leqslant 1]\,\mathrm{d}y\\ &= 0.\end{aligned}$$

所以X,Y不相关. 因为知道X取值x后, Y在

$$[-\sqrt{1-x^2},\sqrt{1-x^2}\,]$$

中取值, X,Y不独立.

§3.7 协方差和相关系数

相关系数 ρ_{XY}只表示了X, Y间的线性关系. 当$\rho_{XY} = 0$, 尽管称X, Y不相关, 它们之间还可以有其他的非线性关系. 例如当$Y = X^2$, $X \sim N(0, 1)$时, (X, Y)的取值总在抛物线$y = x^2$ 上, 但是X, Y不相关. 这是因为X的概率密度$\varphi(x)$是偶函数, $x^3\varphi(x)$是奇函数, 所以

$$\mathrm{E}\, X^3 = \int_{-\infty}^{\infty} x^3 \varphi(x)\,\mathrm{d}x = 0.$$

于是得到

$$\mathrm{Cov}\,(X, Y) = \mathrm{E}\,[X(Y-1)] = \mathrm{E}\, X^3 - \mathrm{E}\, X = 0.$$