概率论复习

第一章：一维变量

一、 概率

1. 随机试验的条件
   1) 可重复性
   2) 不止一个结果，但所有结果范围已知
   3) 实验前发生哪个结果未知
2. **概率计算公式**
   1) **加法公式（偶减奇加）：P(AUB)=P(A)+P(B)-P(AB)**
      P(AUBUC)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)
   2) **乘法公式 P(AB)=P(B)P(A|B)**
   3) **全概率公式（原因推结果）：**
      $$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i)$$
   4) **贝叶斯公式（结果推原因）：**
      $$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n}P(B_j)P(A|B_j)}$$
3. **事件的独立性**
   1) **A、B 独立：P(AB)=P(A)P(B)，P(A)=P(A|B)**
   2) **多事件独立：与上类似；多事件独立→任意两事件独立，反之不成立**
4. **概率分布函数：F(x)=P{X≤x},x∈R，F(x)是 X 的概率分布函数**
   性质：右连续性：F(x)为右连续函数，区间[a,b)，F(x)=F(x+0)

二、 离散变量

1. **退化分布（一点分布）**
2. **两点分布（贝努利分布）：X~B(1,p)。期望：p，方差：p(1-p)**
3. **二项分布（n 重贝努利分布）：X~B(n,p)，$P\{X=k\}=C_n^k p^k(1-p)^{n-k}$,中间项概率最大。**
   期望：np，方差：np(1-p)。X1~B(n1,p), X2~B(n2,p), 则 X1+X2~B(n1+n2,p)
4. **泊松分布（计数）：X~P(λ)，$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},k=0,1,\cdots$，二项分布的极限。期望和方差均为λ。**
   参数 λ 是单位时间(或单位面积)内随机事件的平均发生次数，满足线性相加（2 倍单位时间内…服从 2λ分布）。
   泊松分布适合于描述单位时间内随机事件发生的次数。
5. **几何分布：X~G(p)，$P(X=k)=(1-p)^{k-1}p,k=1,2,\ldots$，期望：$\frac{1}{p}$，方差：$\frac{1-p}{p^2}$**
   意义：n 次伯努利试验中前 k-1 次皆失败，第 k 次成功的概率。

三、 离散变量

1. 概率密度函数 f(x)：$P\{x_1 < X \le x_2\} = \int_{x_1}^{x_2} f(x)dx.$，物理意义：线密度
2. 分布函数 F(x)：F(x)=P{X≤x}, $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$,F'(x)=f(x)
   性质：X 为连续变量→ F(x)为连续函数，反之不成立
3. 均匀分布：X~U[a,b]
   1) 密度函数：$f(x)=\begin{cases}\frac{1}{b-a} & a\le x\le b\\ 0 & \text{其它}\end{cases}$
   2) 分布函数：$F(x)=\frac{x-a}{b-a}, a\le x\le b$
   3) 期望：(a+b)/2，方差：$(b-a)^2/12$
4. 指数分布：X~ E(λ)
   1) 密度函数：$f(x)=\begin{cases}\lambda e^{-\lambda x}, & (x\ge 0)\\ 0, & (x<0)\end{cases}$
   2) 分布函数：$F(x;\lambda)=\begin{cases}1-e^{-\lambda x} & , x\ge 0,\\ 0 & , x<0.\end{cases}$
   3) 期望：$E(X)=\frac{1}{\lambda}$，方差：$D(X)=Var(X)=\frac{1}{\lambda^2}$
   4) 无记忆性：$P(T>s+t|T>t)=P(T>s)$，即，如果 T 是某一元件的寿命，已知元件使用了 t 小时，它总共使用至少 $s+t$ 小时的条件概率，与从开始使用时算起它使用至少 s 小时的概率相等。
5. 正态分布：$X\sim N(\mu,\sigma^2), Y=\frac{X-\mu}{\sigma}\sim N(0,1)$（Y 为标准正态分布）
   1) 密度函数：$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
   2) 期望：μ，方差：σ^2
   3) 如果 $X\sim N(\mu,\sigma^2)$ 且 a 与 b 是实数，那么 $aX+b\sim N(a\mu+b,(a\sigma)^2)$
   4) 如果 $X\sim N(\mu_X,\sigma_X^2)$ 与 $Y\sim N(\mu_Y,\sigma_Y^2)$ 是独立（独立系数ρ=0）的正态随机变量，那么它们的和满足正态分布 $U=X+Y\sim N(\mu_X+\mu_Y,\sigma_X^2+\sigma_Y^2)$，它们的差满足正态分布 $V=X-Y\sim N(\mu_X-\mu_Y,\sigma_X^2+\sigma_Y^2)$.。U 与 V 两者是相互独立的。
   5) 如果 $X_1,\cdots,X_n$ 为独立标准常态随机变量，那么 $X_1^2+\cdots+X_n^2$ 服从自由度为 n 的卡方分布。
   6) 若 Y 满足标准正态分布，则分布|Y|的期望为$\sqrt{\frac{2}{\pi}}$，方差为 1-$\frac{2}{\pi}$

第二章：二维变量

一、　二维变量

1. 二维随机变量(X,Y)：X 与 Y 相互独立
2. (X,Y)的联合分布函数 F(x,y)与联合概率密度函数 f(x,y)：

$$F(x,y)=P(X\le x,Y\le y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(x,y)dxdy.$$

$$f(x,y)=F''_{xy}(x,y);$$

二、　边缘分布与独立性

1. 边缘分布：多维随机变量中只包含其中部分变量的概率分布。如对(X,Y)分布只研究X 的分布。
2. 边缘分布函数函数和边缘概率密度函数：

[X 的边缘分布函数]

$$\begin{aligned}F_X(x)=P(X\le x)&=P(X\le x,Y<+\infty)\\&=F(x,+\infty)\\&=\int_{-\infty}^{x}dx\int_{-\infty}^{+\infty}f(x,y)dy.\end{aligned}$$

[X 的边缘概率密度]

$$f_X(x)=\frac{d}{dx}F_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy.$$

3. 离散变量的独立性：$p(x_i,y_j)=p_X(x_i)p_Y(y_j).$
4. 连续变量的独立性：$F(x,y)=F_X(x)F_Y(y),$ $f(x,y)=f_X(x)f_Y(y).$

第三章：期望、方差、协方差、相关系数

一、 期望 E(X)：X 的平均值

1. 离散型：$E(X)=\sum_{k=1}^{\infty}x_k p_k$

2. 一维连续型：$E(X)=\int_{-\infty}^{\infty}xf(x)dx$

3. 二维连续型 ($Z=g(X,Y)$) ：$E(Z)=E[g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f(x,y)dxdy$

4. 样本均值的期望：$E(\bar{X})=\mu$，μ是样本整体期望

二、 方差 D(X)：随机变量与期望值偏差平方的平均值

1. 表达式（方差=平方的期望-期望的平方）：

$$D(X)=E\left\{[X-E(X)]^2\right\}=E\left\{X^2-2XE(X)+[E(X)]^2\right\}$$

$$=E\left(X^2\right)-2E(X)E(X)+[E(X)]^2=E\left(X^2\right)-[E(X)]^2$$

2. 性质：

   1) $D(CX)=C^2D(X), D(X+C)=D(X)$

   2) $D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$，其中 X 与 Y 独立时 cov(X,Y)=0

   3) 切比雪夫不等式：设随机变量 X 具有数学期望 $E(X)=\mu$，方差 $D(X)=\sigma^2$ 则对任意正数 ε，不等式 $P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}$ 成立。

3. 样本期望的方差：$D(\bar{X})=\frac{\sigma^2}{n}$，$\sigma^2$ 是样本整体方差

   证明：D(ΣXi/n)=ΣD(Xi)/(n^2)=DX/n

4. 统计量的方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$，是总体样本方差的无偏估计

   n-1 而非 n 的原因：$D(\bar{X})\neq 0$，证明如下：

$$E(S^2)=E(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2)=\frac{1}{n-1}E(\sum_{i=1}^{n}X_i^2-n\bar{X}^2)$$

$$=\frac{1}{n-1}(\sum_{i=1}^{n}E(X_i^2)-nE(\bar{X}^2))$$

$$=\frac{1}{n-1}(\sum_{i=1}^{n}[D(X_i)+E^2(X_i)]-n[D(\bar{X})+E^2(\bar{X})])$$

$$=\frac{1}{n-1}(\sum_{i=1}^{n}[\sigma^2+\mu^2]-n[\frac{1}{n}\sigma^2+\mu^2])=\sigma^2$$

三、 协方差 cov(X,Y)=E(XY)-E(X)E(Y)

1. 协方差用于衡量两个变量的总体误差，描述两个变量间的相关程度。而方差是协方差的一种特殊情况，即当两个变量是相同的情况。如果两个变量的变化趋势一致，也就是说如果其中一个大于自身的期望值，另外一个也大于自身的期望值，那么两个变量之间的协方差就是正值。如果两个变量的变化趋势相反，即其中一个大于自身的期望值，另外一个却小于自身的期望值，那么两个变量之间的协方差就是负值。
2. $$\begin{aligned}Cov(X,Y)&=E[(X-E[X])(Y-E[Y])]\\&=E[XY]-2E[Y]E[X]+E[X]E[Y]\\&=E[XY]-E[X]E[Y]\end{aligned}$$
3. 性质
   1) Cov(X，Y)=Cov(Y，X)；
   2) Cov(aX，bY)=abCov(X，Y)，(a，b 是常数)；
   3) Cov(X1+X2，Y)=Cov(X1，Y)+Cov(X2，Y)。

四、 (皮尔逊)相关系数ρ

1. 标准化变量 $X^*=\frac{X-EX}{\sqrt{DX}}$，$E(X^*)=0, D(X^*)=1$
2. 相关系数：ρ(X,Y)=cov(X^*,Y^*)，化简：$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$
3. 相关系数意义：表述 X 和 Y 的线性相关程度
4. 相关系数性质
   1) $|\rho_{XY}|\le 1.$
   2) X 与 Y 独立 ⟶ ρ=0 ⟷ X 与 Y 线性无关
   3) 若 X 与 Y 均为正态分布，则：X 与 Y 独立 ⟷ ρ=0 ⟷ X 与 Y 线性无关
   4) ρ(X,X^2)=0