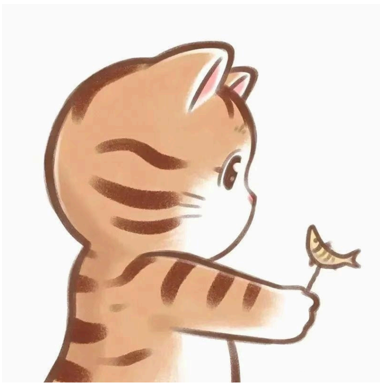
26 春信概统学习笔记

Qiyu Zhang 编著



https://zhangqiyu.me

2025 年 11 月

目录

信息学中的概率统计—— $3'$

〇 · 上课笔记

1.赌资分配问题：两个赌徒进行一系列公平的赌博（每局获胜概率均为1/2），约定先赢得一定局数（如先赢6局）的人可获得全部赌注。然而，比赛因故中断，此时两人已赢得的局数不同（例如甲赢5局，乙赢3局）。问题是如何公平地分配赌注。

解:

- 帕斯卡递归法:

一 · 随机事件与概率

1.1 排列组合

1.加法 & 乘法原理

2.排列 & 组合

1.2 样本空间与随机事件

1.随机现象，事件，统计规律性

2.随机试验需满足以下3条件：
- 可以在相同条件下重复进行
- 每次试验的结果是明确可知的且不止一个
- 每次试验有且仅有一个结果发生，但试验前无法预知

3.样本空间Ω与样本点ω: 随机试验的所有可能结果组成的集合; 样本空间的元素称为样本点。

4.（随机）事件：样本空间的子集

5.基本事件：样本空间的单元素子集。与之相对的是复合事件。

6.随机事件发生：如果试验结果是事件的一个样本点，则称事件发生。

7.必然事件 & 不可能事件

1.3 事件的关系与运算

1.venn 图

2.包含，相等，并（或，和），交（与，积），差，互斥，对立（互斥，二者补且必居其一）

3.完备事件组：若$A_1, A_2, ..., A_n$是样本空间Ω的一个划分，则称$A_1, A_2, ..., A_n$是一个完备事件组。/ 若$A_1, A_2, ..., A_n$两两互斥，且$\Omega = \bigcup\limits_{i=1}^{n} A_i$

显然，对立（互逆）事件构成了一个最小的完备事件组。

4.运算律：交换律，结合律，分配律，德摩根律等

5.运算顺序：逆，交，并，差

1.4 随机事件的概率

1.频率：设在n次独立重复试验中，事件A发生了n_A次，则称$\frac{n_A}{n}$为事件A的频率。

2.频率的性质：非负性，可加性（对于互斥事件），归一性，稳定性

3.随机事件的概率：

- 统计定义：在不变的条件下，重复n次实验，事件A发生了n_A次，并且当n很大时，$\frac{n_A}{n}$稳定地在某一数值p附近摆动，而且一般来说，n越大摆动就越小，则称p为事件A的概率。
- 主观定义：个人对事件发生的信心程度，通常用 0 到 1 之间的数值来表示。

1.5 古典概型与几何概型

1.服从古典概型：基本事件有限且等可能

2.古典概型定义：设Ω是一个有限集合，A是Ω的一个事件，如果Ω中的每个基本事件发生的可能性相等，则称事件A服从古典概型，并且事件A的概率为$p(A) = \frac{|A|}{|\Omega|}$

3.

1.9 几何概型

1.定义：设样本空间为有限区域Ω，样本点为该区域的任意点。若样本点落入Ω内任意子区域A的概率与A的测度成正比，即$P(A) = \frac{S_A}{S_\Omega}$，则称此时的概率模型为几何概型。

【例 1】将一根棍子随机折断成三段，求三段能组成三角形的概率。

解：设棍长为 1，三段长度分别为$x, y, 1-x-y$，则样本空间$\Omega = \{(x,y)|x>0, y>0, x+y<1\}$，事件$A = \{(x,y)|x>0, y>0, 1-x-y>0, x+y>1-x-y, x+1-x-y>y, y+1-x-y>x\}$，因此$P(A) = \frac{S_A}{S_\Omega} = \frac{1}{4}$

【例 2】蒲丰针问题：在平行线上以等距画出多条线段，线段间距为d，将一根长度为$l(l<d)$的针随机抛掷在这些线段上，求针与线段相交的概率。

解：设针的中心点到最近线段的距离为x，针与线段的夹角为θ，则样本空间$\Omega=\left\{(x,\theta)|0\leq x\leq\frac{d}{2},0\leq\theta\leq\frac{\pi}{2}\right\}$，事件$A=\left\{(x,\theta)|0\leq x\leq\frac{l}{2}\sin\theta,0\leq\theta\leq\frac{\pi}{2}\right\}$，因此$P(A)=\frac{S_A}{S_\Omega}=\frac{2l}{\pi d}$

1.10 加法公式

1.加法公式：设$A_1,A_2,...,A_n$是样本空间Ω的n个事件，则

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leq i<j\leq n}P\left(A_i\cap A_j\right)+...+(-1)^{n-1}P\left(\bigcap_{i=1}^{n}A_i\right)$$

【例 1】整除问题（见离散）

【例 2】错排问题（见离散）

1.11 几条事实

1.不可能事件的概率为 0，但概率为 0 的事件不一定是不可能事件。

2.必然事件的概率为 1，但概率为 1 的事件不一定是必然事件。

3.事件A是事件B的子事件，但概率有可能相等。

1.12 概率的公理化定义

1.设Ω为样本空间（非空），记2^Ω为其幂集，$\mathcal{F}\subseteq 2^\Omega$，且
- $\Omega\in\mathcal{F}$
- 若$A\in\mathcal{F}$，则$\overline{A}\in\mathcal{F}$
- 若$A_i\in\mathcal{F}$，则$\bigcup_{i=1}^{\infty}A_i\in\mathcal{F}$

则称$\mathcal{F}$是Ω上的一个σ代数，称$(\Omega,\mathcal{F})$是一个可测空间。

2.设可测空间$(\Omega,\mathcal{F})$，则称$P:\mathcal{F}\to[0,1]$且
- $P(\Omega)=1$
- $0\leq P(A)\leq 1,\forall A\in\mathcal{F}$
- 若$A_i\in\mathcal{F}$且$A_i\cap A_j=0,\forall i\neq j$，则$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$

则称P是$(\Omega,\mathcal{F})$上的一个概率测度，称$(\Omega,\mathcal{F},P)$是一个概率空间，并称$P(A)$为随机事件A的概率。

3.

1.13 条件概率与事件的独立性

1.条件概率定义：设概率空间$(\Omega,\mathcal{F},P)$，对于$B\in\mathcal{F}$且$P(B)>0$，称$P(A|B)=\frac{P(A\cap B)}{P(B)}$为事件$A$在事件$B$发生的条件下的概率。

2.全概率公式：设$A_1,A_2,...,A_n$是样本空间Ω的一个完备事件组，且$P(A_i)>0,\forall i$，则对于任一事件$B\in\mathcal{F}$，有

$$P(B)=\sum_{i=1}^{n}P(B|A_i)P(A_i)$$

3.贝叶斯公式：设$A_1,A_2,...,A_n$是样本空间Ω的一个完备事件组，且$P(A_i)>0,\forall i$，则对于任一事件$B\in\mathcal{F}$且$P(B)>0$，有

$$P(A_i|B)=\frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n}P(B|A_j)P(A_j)}$$

4.先验概率：事件A的概率$P(A)$称为事件A的先验概率。

5.后验概率：事件A在事件B发生的条件下的概率$P(A|B)$称为事件A的后验概率。

6.独立事件：设$A,B\in\mathcal{F}$，如果$P(A\cap B)=P(A)P(B)$，则称事件A与事件B相互独立。

7.三个事件A,B,C相互独立：如果$P(A\cap B)=P(A)P(B),P(A\cap C)=P(A)P(C),P(B\cap C)=P(B)P(C),P(A\cap B\cap C)=P(A)P(B)P(C)$，则称事件$A,B,C$相互独立。

1.14 独立试验序列概型

1.n次独立试验序列：在相同的条件下，重复进行相同的试验n次，每次实验的结果为有限个且各次试验相互独立，这样的n次试验序列称为~

2.伯努利概型：特别地，在独立试验序列概型中，每次试验只有两种可能结果（通常称为成功和失败），这样的独立试验序列概型称为伯努利概型。

3.二项概率公式：在伯努利概型中，设每次试验成功的概率为p，则在n次试验中恰好有k次成功的概率为

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

4.几何分布：在伯努利概型中，设每次试验成功的概率为p，则第一次成功发生在第k次试验的概率为

$$P(X=k)=(1-p)^{k-1}p$$

5.帕斯卡分布：在伯努利概型中，设每次试验成功的概率为p，则第r次成功发生在第k次试验的概率为

$$P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r}$$

二·随机变量及其概率分布

2.1 随机变量

1.随机变量：对概率空间$(\Omega,\mathcal{F},P)$上的一个可测函数$X:\Omega\to R$，若对任意$x\in R$，有$\{\omega\in\Omega:X(\omega)<x\}\in\mathcal{F}$，则称$X$为一个随机变量。

2.类型：离散；非离散

2.2 随机变量的概率分特性刻画

1.分布函数：设X是一个随机变量，则称$F(x)=P(X\le x),-\infty<x<\infty$为随机变量$X$的分布函数。

2.分布函数的性质：

- 非减性：$F(x)$是一个非减函数
- 右连续性：$F(x)$是一个右连续函数
- 极限：$\lim_{x\to-\infty}F(x)=0,\lim_{x\to\infty}F(x)=1$

2.3 离散型随机变量的概率分布

1.分布律：设X是一个离散型随机变量，$x_1,x_2,...$是X的所有可能取值，则称$p_i=P(X=x_i),i=1,2,...$为随机变量X的分布律。常用分布律表示：$X:x_1,x_2,...;P:p_1,p_2,...$

2.性质：

- 非负性：$p_i\ge 0,\forall i$
- 归一性：$\sum_{i=1}^{\infty}p_i=1$

3.分布函数：设X是一个离散型随机变量，$x_1,x_2,...$是X的所有可能取值，则称$F(x)=P(X\le x)=\sum_{x_i\le x}p_i$为随机变量$X$的分布函数。

2.4 离散型随机变量常见分布

1.0-1 分布（两点分布）：设X是一个离散型随机变量，$P(X=1)=p,P(X=0)=1-p$，则称X服从参数为p的 0-1 分布，记为$X\sim B(1,p)$

2.二项分布：设X是一个离散型随机变量，$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},k=0,1,...,n$，则称$X$服从参数为$n,p$的二项分布，记为$X\sim B(n,p)$

3.几何分布：设X是一个离散型随机变量，$P(X=k)=(1-p)^{k-1}p,k=1,2,...$，则称$X$服从参数为$p$的几何分布，记为$X\sim G(p)$

4.帕斯卡分布：设X是一个离散型随机变量，$P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r}, k=r,r+1,...$，则称$X$服从参数为$r,p$的帕斯卡分布，记为$X\sim P(r,p)$

5.超几何分布：设X是一个离散型随机变量，

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, k=\max(0,n-(N-M)),...,\min(n,M)$$

，则称X服从参数为N,M,n的超几何分布，记为$X\sim H(N,M,n)$

6.泊松分布：设X是一个离散型随机变量，$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}, k=0,1,...$，则称$X$服从参数为$\lambda$的泊松分布，记为$X\sim P(\lambda)$

7.泊松定理：设X_n是一个参数为n,p_n的二项分布随机变量，且$\lim_{n\to\infty} np_n=\lambda$，则对于任一非负整数$k$，有$\lim_{n\to\infty} P(X_n=k)=\frac{\lambda^k e^{-\lambda}}{k!}$

证：

$$\begin{aligned}
C_n^k p^k(1-p)^{n-k} &= \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n!}{(n-k)!\times n^k}\times\frac{\lambda^k}{k!}\times\left(1-\frac{\lambda}{n}\right)^n\times\left(1-\frac{\lambda}{n}\right)^{-k} \\
&\to \frac{\lambda^k}{k!}e^{-\lambda}
\end{aligned}$$

8.极限分布：二项分布是超几何分布的极限分布，泊松分布是二项分布的极限分布。

9.泊松分布的应用：稀有事件在大量重复试验中出现的次数服从泊松分布，当n很大且p很小（np是个不大的数）时，二项分布近似于泊松分布。 、

2.5 连续型随机变量及常见分布

定义 2.5.1（概率密度）设X是随机变量，若对于任意一个实数x，存在一个非负可积函数$f(x)$使得

$$F(x)=P(X\le x)=\int_{-\infty}^{x} f(t)dt \qquad -\infty<x<\infty$$

则称X是连续型随机变量，$f(x)$为随机变量X的概率密度。

那么，很容易得出概率密度的性质：

- 非负性：$f(x)\ge 0,\forall x\in R$
- 归一性：$\int_{-\infty}^{\infty} f(x)dx=1$
- $P(a<X\le b)=\int_a^b f(x)dx$
- $f(x)=F'(x)$（如果$F(x)$在x处可导）

而取任一常数的概率为0，因此无需考虑区间的开闭性。也正因此，连续型随机变量的密度不唯一，可以任意改变有限个点上的值，今后不加区别。

下面，我们介绍一些常见的连续型随机变量分布：

定义 2.5.2（均匀分布） 若X概率密度为

$$f(x)=\begin{cases}\frac{1}{b-a} & a<x<b\\ 0 & \text{否则}\end{cases}$$

则称X服从参数为a,b的均匀分布，记为$X\sim U(a,b)$。则X的分布函数为

$$F(x)=\begin{cases}0 & x\le a\\ \frac{x-a}{b-a} & a<x<b\\ 1 & x\ge b\end{cases}$$

定义 2.5.3（指数分布） 若X概率密度为

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x\ge 0\\ 0 & \text{否则}\end{cases}$$

则称X服从参数为$\lambda(\lambda>0)$的指数分布，记为$X\sim E(\lambda)$。则X的分布函数为

$$F(x)=\begin{cases}0 & x<0\\ 1-e^{-\lambda x} & x\ge 0\end{cases}$$

指数分布经常用来描述各种寿命问题，例如电子元件的寿命，人的寿命等。

【例】设某元件使用时间t后，在Δt内失效的概率为$\lambda\Delta t+o(\Delta t)$，$\lambda$为一常数。假定元件寿命为 0 的概率为 0，求元件寿命T的分布函数。

解：

1. 当$t\le 0$，$F(T)=P(T\le 0)=0$

2. 当$t>0$，则

$$\lambda\Delta t+o(\Delta t)=\frac{F(t+\Delta t)-F(t)}{1-F(t)}$$

$$F(t+\Delta t)-F(t)=(1-F(t))(\lambda\Delta t+o(\Delta t))$$

因此

$$F'(t)=\lim_{\Delta t\to 0}\frac{F(t+\Delta t)-F(t)}{\Delta t}=\lambda(1-F(t))$$

则

$$\frac{dF}{dt}=\lambda(1-F)$$

即

$$\frac{dF}{1-F}=\lambda dt$$

则

$$\ln(1-F)=-\lambda t$$

则$F(t)=1-e^{-\lambda t}$

因此

$$X\sim E(\lambda)$$

从上例中，我们也可以看出，指数分布具有无记忆性：对于任一$s,t\geq 0$，有

$$P(X>s+t|X>s)=P(X>t)$$

这是因为

$$\begin{aligned}
P(X>s+t|X>s)&=\frac{P(X>s+t\cap X>s)}{P(X>s)}\\
&=\frac{P(X>s+t)}{P(X>s)}\\
&=\frac{1-F(s+t)}{1-F(s)}\\
&=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}\\
&=e^{-\lambda t}\\
&=P(X>t)
\end{aligned}$$

定义 2.5.4（正态分布）若X概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<\infty$$

则称X服从参数为μ,σ^2的正态分布，记为$X\sim N(\mu,\sigma^2)$。则X的分布函数为

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt,-\infty<x<\infty$$

我们来验证其归一性：令$u=\frac{x-\mu}{\sigma},t=\frac{u^2}{2}$，则

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{u^2}{2}}du\\
&=\frac{1}{\sqrt{\pi}}\int_{0}^{\infty}t^{-\frac{1}{2}}e^{-t}dt\\
&=\frac{1}{\sqrt{\pi}}\times\Gamma\left(\frac{1}{2}\right)\\
&=1
\end{aligned}$$

正态分布的性质无需多言，但有一点需要注意的是，$f(x)$的拐点为$\mu\pm\sigma$。这个结论也有助于我们理解正态分布的形状：当μ固定时，σ越大，分布越平坦。

此外，当$X\sim N(\mu,\sigma^2)$，则$Z=\frac{X-\mu}{\sigma}\sim N(0,1)$，因此可以通过标准正态分布表来计算任一正态分布的概率。此时，记$\varphi(z)$为概率密度，$\Phi(z)=P(Z\leq z)$为标准正态分布的分布函数。

定义 2.5.5（伽马分布）若X概率密度为

$$f(x)=\begin{cases}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} & x\geq 0\\ 0 \ \ 否则\end{cases}$$

则称X服从参数为α,β的伽马分布，记为$X\sim\Gamma(\alpha,\beta)$。

这里补充一下伽马函数：

$$\Gamma(\alpha)=\int_0^\infty t^{\alpha-1}e^{-t}dt,\alpha>0$$

满足

$$\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$$

因此当α为正整数n时，

$$\Gamma(n)=(n-1)!$$

而且

$$\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$$

我们来验证一下伽马分布的归一性：令$u=\beta x$，则

$$\begin{aligned}\int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}dx &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{\alpha-1}e^{-\beta x}dx\\ &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty \left(\frac{u}{\beta}\right)^{\alpha-1}e^{-u}\left(d\frac{u}{\beta}\right)\\ &= \frac{1}{\Gamma(\alpha)}\int_0^\infty u^{\alpha-1}e^{-u}du\\ &= 1\end{aligned}$$

并且当$\alpha=1$时，伽马分布退化为指数分布。即$\Gamma(1,\beta)=E(\beta)$

设n为正整数，$\alpha=\frac{n}{2},\beta=\frac{1}{2}$时，伽马分布退化为卡方分布，记为$X\sim\chi^2(n)$。此时，$X$的概率密度为

$$f(x)=\begin{cases}\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}e^{-\frac{x}{2}} & x\geq 0\\ 0 \ \ 否则\end{cases}$$

2.6 随机变量的函数及其分布

首先讨论离散型随机变量的函数：设X是一个离散型随机变量，$g: R \to R$是一个函数，则$Y = g(X)$也是一个离散型随机变量，且Y的分布律为

$$P(Y = y) = \sum_{x:g(x)=y} P(X = x)$$

至于连续型随机变量，则直接从定义入手或借助已知公式，下面我们以几个例子来说明：

【例】设X密度函数为连续函数$f_X(x)$，求证：$Y = aX + b(a \neq 0)$的密度函数为

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

证：

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right)$$

当$a > 0$时，

$$F_Y(y) = F_X\left(\frac{y-b}{a}\right)$$

当$a < 0$时，

$$F_Y(y) = 1 - F_X\left(\frac{y-b}{a}\right)$$

因此

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

【例】设$X \sim N(\mu, \sigma^2), Y = aX + b(a \neq 0)$，求$Y$的密度函数。

解：由上一例可知，

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) = \frac{1}{|a|\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}$$

因此

$$Y \sim N(a\mu + b, (|a|\sigma)^2)$$

此例也说明了正态随机变量的线性变换仍服从正态分布。

【例】设分子运动速度V服从麦克斯韦分布（参数为α），其密度函数

$$f_X(x)=\begin{cases}\frac{4x^2}{\alpha^3\sqrt{\pi}}e^{-\frac{x^2}{\alpha^2}} & x\geq 0\\ \\ 0 \quad 否则\end{cases}$$

求动能$Y=\frac{1}{2}mV^2$的密度函数。

解：

$$F_Y(y)=P(Y\leq y)=P\left(\frac{1}{2}mV^2\leq y\right)=P\left(V\leq\sqrt{\frac{2y}{m}}\right)$$

当$y<0$时，$F_{Y(y)}=0$ 当$y\geq 0$时，

$$F_Y{}'(y)=f_X\left(\sqrt{\frac{2y}{m}}\right)d\left(\sqrt{\frac{2y}{m}}\right)$$

于是

$$f_Y(y)=\begin{cases}\frac{4\sqrt{2y}}{m^{\frac{3}{2}}\alpha^3\sqrt{\pi}}e^{-\frac{2y}{m\alpha^2}} & y\geq 0\\ \\ 0 \quad 否则\end{cases}$$

三・多维随机变量及其概率分布

3.1 多维随机变量

定义 3.1.1（多维随机变量）设Ω为随机试验的样本空间，若$X_1,X_2,...,X_n$是定义在Ω上的n个随机变量，则称$(X_1,X_2,...,X_n)$为一个n维随机变量。

3.2 多维随机变量的分布函数

定义 3.2.1（多维随机变量的联合分布函数）设$(X_1,X_2,...,X_n)$为一个n维随机变量，则其联合分布函数为$F(x_1,x_2,...,x_n)=P(X_1\leq x_1,X_2\leq x_2,...,X_n\leq x_n)$。

定义 3.2.2（多维随机变量的边缘分布函数）设$(X_1,X_2,...,X_n)$为一个n维随机变量，则对于任意k个索引$i_1,i_2,...,i_k$，可构成一个k维随机变量$\left(X_{i_1},X_{i_2},...,X_{i_k}\right)$，其分布函数称为$(X_1,X_2,...,X_n)$关于其相应$k$个索引的边缘分布函数。

下面我们主要讨论二维的情况。

3.3 二维随机变量

联合分布函数的性质，有

- 取值范围：$0 \le F(x,y) \le 1, F(+\infty,+\infty)=1, F(-\infty,-\infty)=F(x,-\infty)=F(-\infty,y)=0, \forall x,y \in R$
- 非减性
- 右连续

联合分布函数与边缘分布函数的关系：设(X,Y)为一个二维随机变量，则

$$F(x,y)=P(X \le x, Y \le y)$$

因此

$$F_X(x)=P(X \le x)=P(X \le x, Y \le +\infty)=F(x,+\infty)$$

同样地，

$$F_Y(y)=F(+\infty,y)$$

对于二维离散型随机变量(X,Y)，设X的所有可能取值为$x_1,x_2,\dots$，Y的所有可能取值为$y_1,y_2,\dots$，则称$p_{ij}=P\left(X=x_i,Y=y_j\right),i,j=1,2,\dots$为二维随机变量$(X,Y)$的联合分布律。则其联合分布函数为

$$F(x,y)=\sum_{x_i \le x}\sum_{y_j \le y} p_{ij}$$

边缘分布律为

$$P(X=x_i)=\sum_{j=1}^{\infty} p_{ij}, P\left(Y=y_j\right)=\sum_{i=1}^{\infty} p_{ij}$$

并称p_{ij}为联合概率分布

而边缘概率分布

$$p_{i\cdot}=\sum_{j=1}^{\infty} p_{ij}$$

$$p_{\cdot j}=\sum_{i=1}^{\infty} p_{ij}$$

可见，由二维离散型随机变量的联合概率分布，可以唯一确定边缘概率分布（通过对另一个变量求和即可得到）。但反之不成立。

但是二维离散型随机变量的联合分布函数和联合分布律可以相互求出：

$$F(x,y)=\sum_{x_i \le x}\sum_{y_j \le y} p_{ij}$$

$$p_{ij}=F\left(x_i,y_j\right)-F\left(x_i-0,y_j\right)-F\left(x_i,y_j-0\right)+F\left(x_i-0,y_j-0\right)$$

下面我们讨论二维连续型随机变量：

定义 3.3.1（联合概率密度函数）设(X,Y)为一个二维随机变量，若存在非负可积函数$f(x,y)$，使得对于任意x,y，有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dvdu$$

则称$f(x,y)$为(X,Y)的联合概率密度函数。

由高数下二重积分知识，我们有这样一条结论：若$f(x,y)$在点(x,y)处连续，则

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dvdu$$

在点(x,y)处可导，且

$$f(x,y)=\frac{\partial^2 F}{\partial x\partial y}$$

最后，我们来研究二维随机变量的常见分布：

定义 3.3.2（二维均匀分布）设G是平面上的有界区域，面积为A，若(X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}\frac{1}{A} & (x,y)\in G\\ 0 & \text{否则}\end{cases}$$

记作$(X,Y)\sim U(G)$，称(X,Y)服从二维均匀分布。

并且边平行于坐标轴的矩形域上的均匀分布，其边缘分布仍为均匀分布（一维）。

【例】$(X,Y)\sim U(D),D=\{(x,y):0<x<1,0<y<x^2\}$，求$(X,Y)$的边缘概率密度。

解：$|D|=\int_0^1 x^2dx=\frac{1}{3}$，则联合概率密度$f(x,y)=\begin{cases}3 & (x,y)\in D\\ 0 & \text{否则}\end{cases}$，那么

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\begin{cases}0 & x\notin D\\ 3x^2 & x\in D\end{cases}$$

同理，

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\begin{cases}0 & y\notin D\\ 3(1-\sqrt{y}) & y\in D\end{cases}$$

定义 3.3.3（二维正态分布）设(X,Y)为一个二维随机变量，若其联合概率密度函数为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right)}$$

则称(X,Y)服从二维正态分布，记作$(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$。

二维正态分布的边缘分布仍为正态分布：

$$X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$$

并且当$\rho=0$时，二维正态分布退化为两个独立的一维正态分布。

【例】设$(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$，求边缘概率密度。

解:

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

同理,

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

若推至n维正态分布$N(\mu,\Sigma)$，n维随机变量$(X_1,X_2,\ldots,X_n)$的联合概率密度函数为

$$f(x_1,x_2,\ldots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\Sigma)}}e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

其中μ为n维均值向量，Σ为n维协方差矩阵。

对 2 维正态分布，$\mu=(\mu_1,\mu_2)^T,\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix},|\Sigma|=\sigma_1^2\sigma_2^2(1-\rho^2)$

3.4 随机变量的相互独立性

定义 3.4.1（二维随机变量的相互独立性）设(X,Y)为一个二维随机变量，其联合分布函数为$F(x,y)$，边缘分布函数分别为$F_X(x)$和$F_Y(y)$，若对于任意$x,y\in R$，都有$F(x,y)=F_X(x)F_Y(y)$，则称X和Y相互独立。

【例】二维连续随机变量(X,Y)相互独立$\Leftrightarrow f(x,y)=f_X(x)f_Y(y)$（连续点处）

证:

$\Leftarrow$: $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dvdu=\int_{-\infty}^{x}f_X(u)du\int_{-\infty}^{y}f_Y(v)dv=F_X(x)F_Y(y)$

$\Rightarrow$: $F(x,y)=F_X(x)F_Y(y)$，则

$$f(x,y)=\frac{\partial^2F}{\partial x\partial y}=\frac{\partial F_X}{\partial x}\frac{\partial F_Y}{\partial y}=f_X(x)f_Y(y)$$

【例】设$(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$，则$X$和$Y$相互独立$\Leftrightarrow\ \rho=0$

证:

$\Leftarrow$:$\rho=0$, 则$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right)}=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}=f_X(x)f_Y(y)$

$\Rightarrow$:X和Y相互独立，则

$$f(x,y)=f_X(x)f_Y(y)=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

因此

$$\rho=0$$

【例】设X,Y为相互独立的随机变量，$u(x),v(y)$为连续函数，则$u(X),v(Y)$也相互独立。

证：

$$\begin{aligned}
F(u(X),v(Y)) &= P(u(X)\le u(x),v(Y)\le v(y)) \\
&= P(X\le x,Y\le y) \\
&= F(x,y) \\
&= F_X(x)F_Y(y) \\
&= F_{u(X)}(u(x))F_{v(Y)}(v(y))
\end{aligned}$$

下面，我们有两个判断二维连续随机变量相互独立的常用方法：

定理 **3.4.1**（判断二维连续随机变量相互独立）设(X,Y)为二维连续随机变量，其联合概率密度函数为$f(x,y)$，$r(x),g(y)$为非负可积函数，且$f(x,y)=r(x)g(y),\forall x,y\in R$则$X$和$Y$相互独立。且有

$$f_X(x)=\frac{r(x)}{\int_{-\infty}^{+\infty}r(x)dx}$$

$$f_Y(y)=\frac{g(y)}{\int_{-\infty}^{+\infty}g(y)dy}$$

定理 **3.4.2**（判断二维连续随机变量相互独立）设(X,Y)为二维连续随机变量，其联合分布函数为$F(x,y)$，则X和Y相互独立$\Leftrightarrow F(x,y)=G(x)H(y)$，且有

$$F_X(x)=\frac{G(x)}{G(+\infty)},F_Y(y)=\frac{H(y)}{H(+\infty)}$$

上述定义和两大结论也可以推广到n维随机变量的情况，只需视$X=(X_1,X_2,...,X_n),Y=(Y_1,Y_2,...,Y_m)$即可。

3.5 随机变量的条件分布

老规矩，我们先从二维离散型随机变量的条件分布开始：

定义 **3.5.1**（二维离散型随机变量的条件概率分布）设(X,Y)为一个二维随机变量，

$$P\big(Y=y_j\mid X=x_i\big)=\frac{p_{ij}}{p_{i\cdot}}$$

为在$X=x_i$条件下$Y=y_j$的条件概率，则称$P\big(Y=y_j\mid X=x_i\big)$为$(X,Y)$的条件概率分布。

定理 **3.5.2**（乘法公式）

- **(1)**$P(AB)=P(A)P(B|A)=P(B)P(A|B)$

- (2) $P(A_1A_2...A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)...P(A_n|A_1...A_{n-1})$

定理 **3.5.3**（全概率公式）设$B_1,B_2,...,B_n$为事件A的一个划分，则

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

定理 **3.5.4**（贝叶斯公式）设$B_1,B_2,...,B_n$为事件A的一个划分，则对于任一$i=1,2,...,n$，有

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

【例】设射手命中率为p，射击至射中两次为止，记X为第一次射中时的射击次数，Y为第二次射中时的射击次数，分别求X,Y的条件概率分布。

解:

$$p_{ij}=p(1-p)^{i-1}\times p(1-p)^{j-i-1}=p^2(1-p)^{j-2}$$

$$p_{i\cdot}=p(1-p)^{i-1}$$

$$p_{\cdot j}=\sum_{i=1}^{j-1}p_{ij}=(j-1)p^2(1-p)^{j-2}$$

则

$$P(Y=y_j\mid X=x_i)=\frac{p_{ij}}{p_{i\cdot}}=\frac{p(1-p)^{j-2}}{p(1-p)^{i-1}}=p(1-p)^{j-i-1}$$

$$P(X=x_i\mid Y=y_j)=\frac{p_{ij}}{p_{\cdot j}}=\frac{p(1-p)^{j-2}}{\sum_{i=1}^{j-1}p_{ij}}=\frac{1}{j-1}$$

那对于一般的二维连续型随机变量，我们可以直接从定义入手:

定义 **3.5.5**（二维连续型随机变量的条件概率密度函数）设(X,Y)为一个二维随机变量，$f(x,y)$为其联合概率密度函数，$f_X(x),f_Y(y)$为其边缘概率密度函数，则在$X=x$条件下Y的条件概率密度函数为

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}$$

在$Y=y$条件下X的条件概率密度函数为

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$$

有类似的乘法公式、全概率公式和贝叶斯公式:

- (1) $f(x,y)=f_X(x)f_{Y|X}(y|x)=f_Y(y)f_{X|Y}(x|y)$
- (2) $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\int_{-\infty}^{+\infty}f_Y(y)f_{X|Y}(x|y)dy$
- (3) $f_{X|Y}(x|y)=\frac{f_Y(y)f_{X|Y}(x|y)}{\int_{-\infty}^{+\infty}f_Y(y)f_{X|Y}(x|y)dy}$

【例】设$(X,Y)\sim U(D),D=\{(x,y):0<x<1,0<y<x^2\}$，求$f_{X|Y}(x|y),f_{Y|X}(y|x)$。

解:

$$f(x,y)=\begin{cases}3 & (x,y)\in D\\0 & 否则\end{cases}$$

则

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\begin{cases}\frac{1}{1-\sqrt{y}} & y\in D\\0 & 否则\end{cases}$$

同理,

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\begin{cases}\frac{1}{x^2} & x\in D\\0 & 否则\end{cases}$$

【例】设$X\sim U(0,1)$，又设$X=x$时，$Y\sim U(0,x)$，求X,Y的联合概率密度。

解: $f_X(x)=\begin{cases}1 & 0<x<1\\0 & 否则\end{cases}$，$f_{Y|X}(y|x)=\begin{cases}\frac{1}{x} & 0<y<x\\0 & 否则\end{cases}$则

$$f(x,y)=f_X(x)f_{Y|X}(y|x)=\begin{cases}\frac{1}{x} & 0<y<x<1\\0 & 否则\end{cases}$$

当然，上述条件概率相关概念也可以推广至n维。

3.6 二维随机变量函数的分布

【问题】已知二维随机变量(X,Y)的分布，求$Z=g(X,Y)$的分布。

解:

1.当(X,Y)离散时，设X的所有可能取值为$x_1,x_2,\ldots$，Y的所有可能取值为$y_1,y_2,\ldots$，则$Z=g(X,Y)$的分布律为

$$P(Z=z)=\sum_{g(x_i,y_j)=z}P\big(X=x_i,Y=y_j\big)$$

2.当(X,Y)连续时，设$f(x,y)$为(X,Y)的联合概率密度函数，则$Z=g(X,Y)$的分布函数为

$$F(z)=P(Z\le z)=P(g(X,Y)\le z)=\iint_{D_Z}f(x,y)dxdy$$

其中$D_Z=\{(x,y)\mid g(x,y)\le z\}$

下面介绍几类常见二维随机变量函数的分布:

两个具有可加性的离散分布：设$X_1, X_2, ..., X_n$为n个相互独立的随机变量，且$X_i \sim B(m_i, p), i = 1, 2, ..., n$，则$Y = X_1 + X_2 + ... + X_n \sim B\left(\sum_{i=1}^{n} m_i, p\right)$；设$X_1, X_2, ..., X_n$为$n$个相互独立的随机变量，且$X_i \sim P(\lambda_i), i = 1, 2, ..., n$，则$Y = X_1 + X_2 + ... + X_n \sim P\left(\sum_{i=1}^{n} \lambda_i\right)$

对于一般情况，有如下推导：对于$Z = X + Y$

$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z) = \iint_{D_Z} f(x, y) dx dy$$

其中$D_Z = \{(x, y) \mid x + y \leq z\}$，则

$$F_Z(z) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} f(x, y) dy$$

令$u = y + x$则

$$F_Z(z) = \int_{-\infty}^{z} du \int_{-\infty}^{+\infty} f(x, u - x) dx$$

则

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z - x) dx = \int_{-\infty}^{+\infty} f(z - y, y) dy$$

特别地，若X, Y相互独立，则

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z - x) dx = \int_{-\infty}^{+\infty} f_X(z - y) f_Y(y) dy$$

这就是卷积公式。

利用卷积公式，事实上还能证出：设$X_1, X_2, ..., X_n$为n个相互独立的随机变量，且$X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, ..., n$，则$Y = X_1 + X_2 + ... + X_n \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$。还有$(X, Y) \sim N(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho)$，则$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)$

【例】设X, Y相互独立，且都服从$E(\lambda)$，求$Z = X + Y$的密度。

解：

$$f_Z(z) = \int_0^z f_X(x) f_Y(z - x) dx = \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)} dx = \lambda^2 e^{-\lambda z} \int_0^z dx = \lambda^2 z e^{-\lambda z}$$

则

$$f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z} & z > 0 \\ 0 \ \text{否则} \end{cases}$$

若对$Z = aX + bY + c$，有

$$f_Z(z) = \frac{1}{|b|} \int_{-\infty}^{+\infty} f\left(x, \frac{z - ax - c}{b}\right) dx = \frac{1}{|a|} \int_{-\infty}^{+\infty} f\left(\frac{z - by - c}{a}, y\right) dy$$

至于另一种计算$f_Z(z)$的方法，构造新二维随机变量(Z,U)，解出边缘密度。已知$\begin{cases} z=g(x,y) \\ u=r(x,y) \end{cases}$存在唯一反函数$\begin{cases} x=h(z,u) \\ y=s(z,u) \end{cases}$且$h,s$有连续偏导，记

$$J(z,u)=\begin{vmatrix} h_z & h_u \\ s_z & s_u \end{vmatrix}$$

则

$$f_{ZU}(z,u)=f_{XY}(h(z,u),s(z,u))|J|$$

下面讨论商的分布$Z=\frac{X}{Y}$，令$\begin{cases} z=\frac{x}{y} \\ u=y \end{cases}$，则$\begin{cases} x=zu \\ y=u \end{cases}$，则$|J|=\det\left(\begin{vmatrix} z_u & z_u \\ s_u & s_u \end{vmatrix}\right)=|u|$则

$$f_Z(z)=\int_{-\infty}^{+\infty} f_{ZU}(z,u)du=\int_{-\infty}^{+\infty} f(zu,u)|u|du$$

平方和的分布$Z=X^2+Y^2$，则$|J|=\det\left(\begin{vmatrix} s_z & s_u \\ s_z & s_u \end{vmatrix}\right)=\frac{1}{2}$则

$$f_Z(z)=\frac{1}{2}\int_0^{2\pi} f\left(\sqrt{z}\cos u,\sqrt{z}\sin u\right)du$$

相互独立的随机变量的极值分布，对于离散型，可直接计算；对于连续型，设X,Y相互独立，且$F_X(x),F_Y(y)$为其边缘分布函数，则

$$F_{\max(X,Y)}(z)=P(\max(X,Y)\le z)=P(X\le z,Y\le z)=F_X(z)F_Y(z)$$

同样地，

$$F_{\min(X,Y)}(z)=P(\min(X,Y)\le z)=1-P(X>z,Y>z)=1-(1-F_X(z))(1-F_Y(z))$$

四・随机变量的数字特征

4.1 数学期望

定义 4.1.1（离散型随机变量的数学期望）设X为离散随机变量，其概率分布为$P(X=x_i)=p_i$，若无穷级数$\sum_i |x_i|\ p_i$绝对收敛（以防止交换次序后级数值改变），则X的数学期望为$E(X)=\sum_i x_i p_i$。

定义 4.1.2（连续型随机变量的数学期望）设X为连续随机变量，其概率密度函数为$f(x)$，若积分$\int_{-\infty}^{+\infty} xf(x)dx$绝对收敛，则$X$的数学期望为$E(X)=\int_{-\infty}^{+\infty} xf(x)dx$。

定义 4.1.3（一般情况下的数学期望）对随机变量X，任给$\varepsilon>0$，构造离散随机变量$X^\star$：将X所有取值即数轴划分为长为ε的小区间，并定义新随机变量$X^\star=k\varepsilon$，当$X\in[k\varepsilon,(k+1)\varepsilon)$时。故当$E(X^\star),\lim_{\varepsilon\to 0} E(X^\star)$存在时，可定义$E(X)=\lim_{\varepsilon\to 0} E(X^\star)$

4.2 常见分布及随机变量函数的期望

首先讨论离散型随机变量的常见分布：

定理 **4.2.1**（二项分布的数学期望）设$X \sim B(n,p)$，则$E(X)=np$。

证：

$$E(X)=\sum_{k=0}^{n} kC_n^k p^k (1-p)^{n-k}$$

令

$$kC_n^k = nC_{n-1}^{k-1}$$

则

$$E(X)=n\sum_{k=1}^{n} C_{n-1}^{k-1} p^k (1-p)^{n-k}$$

令

$$j=k-1$$

则

$$E(X)=np\sum_{j=0}^{n-1} C_{n-1}^{j} p^j (1-p)^{(n-1)-j}=np$$

定理 **4.2.2**（几何分布的数学期望）设$X \sim G(p)$，则$E(X)=\frac{1}{p}$。

证：

$$E(X)=\sum_{k=1}^{\infty} kp(1-p)^{k-1}$$

令

$$j=k-1$$

则

$$E(X)=p\sum_{j=0}^{\infty}(j+1)(1-p)^j$$

又

$$\sum_{j=0}^{\infty}(j+1)x^j=\frac{1}{(1-x)^2}, |x|<1$$

因此

$$E(X)=p\frac{1}{(1-(1-p))^2}=\frac{1}{p}$$

定理 **4.2.3**（超几何分布的数学期望）设$X \sim H(N,M,n)$，则$E(X)=n\frac{M}{N}$。

证：

$$E(X)=\sum_{k=\max\{0,n-(N-M)\}}^{\min\{n,M\}}k\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}$$

令

$$j=k-1$$

则

$$\begin{aligned}E(X)&=\frac{M}{N}\sum_{j=\max\{-1,n-(N-M)-1\}}^{\min\{n-1,M-1\}}(j+1)\frac{C_{M-1}^jC_{N-M}^{n-1-j}}{C_{N-1}^{n-1}}\\&=\frac{M}{N}\sum_{j=0}^{n-1}(j+1)\frac{C_{M-1}^jC_{N-M}^{n-1-j}}{C_{N-1}^{n-1}}\\&=\frac{M}{N}n\end{aligned}$$

定理 4.2.4（泊松分布的数学期望） 设$X\sim P(\lambda)$，则$E(X)=\lambda$。

证：

$$E(X)=\sum_{k=0}^{\infty}k\frac{\lambda^ke^{-\lambda}}{k!}$$

令

$$j=k-1$$

则

$$E(X)=\lambda\sum_{j=0}^{\infty}\frac{\lambda^je^{-\lambda}}{j!}=\lambda$$

下面我们来讨论连续型随机变量的常见分布：

定理 4.2.5（均匀分布的数学期望） 设$X\sim U(a,b)$，则$E(X)=\frac{a+b}{2}$。 证：

$$E(X)=\int_a^b x\frac{1}{b-a}dx=\frac{a+b}{2}$$

定理 4.2.6（指数分布的数学期望） 设$X\sim E(\lambda)$，则$E(X)=\frac{1}{\lambda}$。

证：

$$E(X)=\int_0^{\infty}x\lambda e^{-\lambda x}dx$$

令

$$t = \lambda x$$

则

$$E(X) = \frac{1}{\lambda}\int_0^{\infty} te^{-t}dt = \frac{1}{\lambda}$$

定理 **4.2.7**（正态分布的数学期望）设$X \sim N(\mu, \sigma^2)$，则$E(X) = \mu$。

证:

$$E(X) = \int_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

令

$$t = \frac{x-\mu}{\sigma}$$

则

$$E(X) = \mu + \sigma\int_{-\infty}^{+\infty} t\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt = \mu$$

定理 **4.2.8**（伽马分布的数学期望）设$X \sim \Gamma(\alpha, \beta)$，则$E(X) = \frac{\alpha}{\beta}$。

证:

$$E(X) = \int_0^{\infty} x\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}dx$$

令

$$t = \beta x$$

则

$$E(X) = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}\frac{1}{\beta} = \frac{\alpha}{\beta}$$

对于随机变量函数的期望，我们有以下结论:

定理 **4.2.9**（随机变量函数的数学期望）设X为随机变量，$g: R \to R$为一个函数，则$Y = g(X)$的数学期望为$E(Y) = E(g(X)) = \sum_i g(x_i)p_i$（离散型）或$E(Y) = E(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$（连续型）。

定理 **4.2.10**（二维随机变量函数的数学期望）设(X,Y)为二维随机变量，$g: R^2 \to R$为一个函数，则$Z = g(X,Y)$的数学期望为$E(Z) = E(g(X,Y)) = \sum_i\sum_j g(x_i, y_j)p_{ij}$（离散型）或$E(Z) = E(g(X,Y)) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dydx$（连续型）。

上述两条定理也可以推广到n维随机变量的情况，只需视$X = (X_1, X_2, ..., X_n), g: R^n \to R$即可。

4.3 数学期望的性质

数学期望的性质有：

- 线性性质：

$$E(aX+b)=aE(X)+b, E(X+Y)=E(X)+E(Y)$$

- X,Y相互独立$\Rightarrow E(XY)=E(X)E(Y)$(反之不成立，可推广到$n$维随机变量的情况)
- 斯瓦尔茨不等式：

$$[E(XY)]^2 \le E(X^2)E(Y^2)$$

下面简略给出施瓦尔茨不等式的证明：

$$E((X-E(X))(Y-E(Y)))^2 \ge 0$$

展开后可得

$$E(XY)^2 \le E(X^2)E(Y^2)$$

4.4 方差及常见分布的方差

定义 4.4.1（方差）设X为随机变量，若$E(X^2)$存在，则称X的方差为$D(X)/Var(X)=E((X-E(X))^2)=E(X^2)-E(X)^2$。称$\sigma=\sqrt{D(X)}$为$X$的标准差或均方差。

常见分布的方差如下：

定理 4.4.2（二项分布的方差）设$X\sim B(n,p)$，则$D(X)=np(1-p)$。

证：

$$D(X)=E(X^2)-E(X)^2$$

又

$$E(X^2)=\sum_{k=0}^{n}k^2C_n^kp^k(1-p)^{n-k}$$

令

$$k(k-1)C_n^k=n(n-1)C_{n-2}^{k-2}$$

则

$$E(X^2)=n(n-1)p^2\sum_{j=0}^{n-2}C_{n-2}^{j}p^j(1-p)^{(n-2)-j}+np\sum_{j=0}^{n-1}C_{n-1}^{j}p^j(1-p)^{(n-1)-j}=n(n-1)p^2+np$$

因此

$$D(X)=np(1-p)$$

定理 4.4.3（几何分布的方差）设$X \sim G(p)$，则$D(X) = \frac{1-p}{p^2}$。

证：

$$D(X) = E(X^2) - E(X)^2$$

又

$$E(X^2) = \sum_{k=1}^{\infty} k^2 p(1-p)^{k-1}$$

令

$$j = k - 1$$

则

$$E(X^2) = p\sum_{j=0}^{\infty} (j+1)^2(1-p)^j$$

又

$$\sum_{j=0}^{\infty} (j+1)^2 x^j = \frac{1+x}{(1-x)^3}, |x| < 1$$

因此

$$E(X^2) = \frac{2-p}{p^2}$$

因此

$$D(X) = \frac{1-p}{p^2}$$

定理 4.4.4（超几何分布的方差）设$X \sim H(N, M, n)$，则$D(X) = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}$。

证：

$$D(X) = E(X^2) - E(X)^2$$

又

$$E(X^2) = \sum_{k=\max\{0,n-(N-M)\}}^{\min\{n,M\}} k^2 \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

令

$$j = k - 1$$

则

$$
\begin{aligned}
E(X^2) &= \frac{M}{N}\sum_{j=\max\{-1,n-(N-M)-1\}}^{\min\{n-1,M-1\}}(j+1)^2\frac{C_{M-1}^{j}C_{N-M}^{n-1-j}}{C_{N-1}^{n-1}} \\
&= \frac{M}{N}\sum_{j=0}^{n-1}(j+1)^2\frac{C_{M-1}^{j}C_{N-M}^{n-1-j}}{C_{N-1}^{n-1}} \\
&= \frac{M}{N}\left(\frac{(M-1)(N-M)}{N(N-1)}n(n-1)+\frac{M}{N}n\right) \\
&= \frac{M}{N}\left(\frac{(M-1)(N-M)}{N(N-1)}n(n-1)+\frac{M}{N}n\right)
\end{aligned}
$$

因此

$$
D(X) = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}
$$

定理 4.4.5（泊松分布的方差）设$X \sim P(\lambda)$，则$D(X)=\lambda$。

证：

$$
D(X) = E(X^2) - E(X)^2
$$

又

$$
E(X^2) = \sum_{k=0}^{\infty}k^2\frac{\lambda^k e^{-\lambda}}{k!}
$$

令

$$
j = k-1
$$

则

$$
E(X^2) = \lambda\sum_{j=0}^{\infty}(j+1)\frac{\lambda^j e^{-\lambda}}{j!} = \lambda(\lambda+1)
$$

因此

$$
D(X) = \lambda
$$

定理 4.4.6（均匀分布的方差）设$X \sim U(a,b)$，则$D(X)=\frac{(b-a)^2}{12}$。

证：

$$
D(X) = E(X^2) - E(X)^2
$$

又

$$
E(X^2) = \int_a^b x^2\frac{1}{b-a}dx = \frac{a^2+ab+b^2}{3}
$$

因此

$$D(X) = \frac{(b-a)^2}{12}$$

定理 **4.4.7**（指数分布的方差）设$X \sim E(\lambda)$，则$D(X) = \frac{1}{\lambda^2}$。

证：

$$D(X) = E(X^2) - E(X)^2$$

又

$$E(X^2) = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx$$

令

$$t = \lambda x$$

则

$$E(X^2) = \frac{1}{\lambda^2} \int_0^{\infty} t^2 e^{-t} dt = \frac{2}{\lambda^2}$$

因此

$$D(X) = \frac{1}{\lambda^2}$$

定理 **4.4.8**（正态分布的方差）设$X \sim N(\mu, \sigma^2)$，则$D(X) = \sigma^2$。

证：

$$D(X) = E(X^2) - E(X)^2$$

又

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

令

$$t = \frac{x-\mu}{\sigma}$$

则

$$E(X^2) = \sigma^2 \int_{-\infty}^{+\infty} t^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt + \mu^2 = \mu^2 + \sigma^2$$

因此

$$D(X) = \sigma^2$$

定理 **4.4.9**（伽马分布的方差）设$X \sim \Gamma(\alpha, \beta)$，则$D(X) = \frac{\alpha}{\beta^2}$。

证:

$$D(X) = E(X^2) - E(X)^2$$

又

$$E(X^2) = \int_0^\infty x^2 \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx$$

令

$$t = \beta x$$

则

$$E(X^2) = \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{1}{\beta^2}$$

因此

$$D(X) = \frac{\alpha}{\beta^2}$$

4.5 方差的性质

方差的性质有:

- $D(aX + b) = a^2 D(X)$
- X, Y相互独立$\Rightarrow D(X + Y) = D(X) + D(Y)$（反之不成立，可推广到$n$维随机变量的情况）
- $D(X \pm Y) = D(X) + D(Y) \pm 2E((X - E(X))(Y - E(Y))) = 2Cov(X, Y)$
- $D(X) \leq E(X - C)^2$当且仅当$C = E(X)$时等号成立。
- $D(X) = 0 \Leftrightarrow P(X = E(X)) = 1$

此外，还有大名鼎鼎的切比雪夫不等式:

定理 **4.5.1**（切比雪夫不等式）设X为随机变量，$\varepsilon > 0$，则

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}$$

证:

$$P(|X - E(X)| \geq \varepsilon) = P((X - E(X))^2 \geq \varepsilon^2)$$

又

$$P((X - E(X))^2 \geq \varepsilon^2) \leq \frac{E((X - E(X))^2)}{\varepsilon^2} = \frac{D(X)}{\varepsilon^2}$$

4.6 其他典型数字特征

定义 4.6.1（偏度系数）设X为随机变量，若$E((X-E(X))^3)$存在，则称$\alpha = \frac{E((X-E(X))^3)}{\sigma^3}$为$X$的偏度系数。

用以刻画随机变量取值关于其数学期望对称程度，易见，当X取值关于$E(X)$对称时，$\alpha = 0$。

定义 4.6.2（峰度系数）设X为随机变量，若$E((X-E(X))^4)$存在，则称$\gamma = \frac{E((X-E(X))^4)}{\sigma^4}$为$X$的峰度系数。

用以刻画随机变量在期望和方差确定时，其概率分布的峰态。如对连续型随机变量，刻画其密度函数曲线的陡峭状态。易证，当X服从正态分布时，$\gamma = 3$。

定义 4.6.3（变异系数）设X为非负随机变量，且$E(X) > 0$，则称$\nu = \frac{\sigma}{E(X)}$为$X$的变异系数。

与方差类似，变异系数也是用以刻画随机变量取值的离散程度，但与方差不同的是，变异系数是一个无量纲的指标，可以用来比较不同随机变量的离散程度。

下面介绍我们熟知的分位数和中位数：

定义 4.6.4（分位数）设X为随机变量，$0 < p < 1$，则称$P(X < a) \le p \le P(X \le a)$为$X$的$p$分位数。

p分位数一定存在，但不一定唯一。

定义 4.6.5（中位数）设X为随机变量，若存在$x \in R$使得$F(x) \ge \frac{1}{2}$且$F(x-0) \le \frac{1}{2}$，则称x为X的中位数。

4.7 协方差与相关系数

对于两个随机变量X, Y，除了独立性之外，我们还可以用协方差和相关系数来刻画它们之间的关系：

定义 4.7.1（协方差）设X, Y为两个随机变量，若$E((X-E(X))(Y-E(Y)))$存在，则称$Cov(X,Y) = E((X-E(X))(Y-E(Y)))$为$X, Y$的协方差。

对于协方差，我们有以下性质：

- $Cov(X,Y) = E(XY) - E(X)E(Y)$
- X, Y相互独立$\Rightarrow Cov(X,Y) = 0$
- $Cov(X,X) = D(X)$
- $Cov(a_1X + b_1, a_2Y + b_2) = a_1a_2Cov(X,Y)$
- $Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$
- $D(X \pm Y) = D(X) + D(Y) \pm 2Cov(X,Y)$
- $Cov(X,Y)^2 \le D(X)D(Y)$（同样是施瓦尔茨不等式，其实就是$\vec{a} \cdot \vec{b} \le \|\vec{a}\| \cdot \|\vec{b}\|$）

随机变量分布可作标准化：可令

$$X^{\star} = \frac{X - E(X)}{\sigma}$$

使得$E(X^{\star})=0,D(X^{\star})=1$，则$X^{\star}$称为$X$的标准化随机变量。

定义 4.7.2（相关系数） 设X,Y为两个随机变量，若$D(X)>0,D(Y)>0$，则称$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}=Cov(X^{\star},Y^{\star})$为$X,Y$的相关系数。

相关系数的性质有：

- $-1\leq\rho_{XY}\leq 1$
- $\rho_{XY}=1$当且仅当$Y=aX+b,a>0$；$\rho_{XY}=-1$当且仅当$Y=aX+b,a<0$。
- X,Y相互独立$\Rightarrow\rho_{XY}=0$（反之不成立）但是X,Y相互独立$\Leftrightarrow\rho_{XY}=0$对于二维正态分布的情况成立。

【例】设$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求$Cov(X,Y)$和$\rho_{XY}$

解：令$u=\frac{X-E(X)}{\sigma_1}$，$v=\frac{Y-E(Y)}{\sigma_2}$，则

$$Cov(X,Y)=E((X-E(X))(Y-E(Y)))=\sigma_1\sigma_2E(uv)=\sigma_1\sigma_2\rho$$

因此

$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}=\rho$$

【例】设(X,Y)服从单位圆上的均匀分布，即$f(x,y)=\begin{cases}\frac{1}{\pi} & x^2+y^2\leq 1\\ 0 & else\end{cases}$求$Cov(X,Y)$并问$X,Y$是否独立。

解：

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$

又

$$E(X)=\iint_{x^2+y^2\leq 1}x\frac{1}{\pi}dxdy=0$$

$$E(Y)=\iint_{x^2+y^2\leq 1}y\frac{1}{\pi}dxdy=0$$

$$E(XY)=\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}}xy\frac{1}{\pi}dydx=0$$

因此

$$Cov(X,Y)=0$$

但X,Y不独立，因为

$$f(x,y)\neq f_X(x)f_Y(y)$$

因为$f_X(x)=\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}}f(x,y)dy=\begin{cases}\frac{2\sqrt{1-x^2}}{\pi} & |x|\leq 1\\ 0 & else\end{cases}$因此$f_X(x)f_Y(y)=\begin{cases}\frac{4\sqrt{(1-x^2)(1-y^2)}}{\pi^2} & |x|\leq 1\\ |y|\leq 1 & \\ 0 & else\end{cases}$因此$X,Y$不独立。

4.8 随机变量的矩

定义 4.8.1（随机变量的矩）设X为随机变量，$r \in N^+$，则

- $E(X^k)$称为X的k阶原点矩
- $E(|X^k|)$称为X的k阶绝对原点矩
- $E((X - E(X))^k)$称为X的k阶中心矩
- $E(X^k Y^l)$称为X, Y的$k + l$阶混合原点矩

定义 4.8.2（n维随机变量的协方差矩阵）设$X = (X_1, X_2, ..., X_n)$为n维随机变量，若对于任意$i, j = 1, 2, ..., n$，$Cov(X_i, X_j)$存在，则称$\Sigma = \left(Cov(X_i, X_j)\right)_{n \times n}$为$X$的协方差矩阵。

定理 4.8.1（协方差矩阵的性质）设$X = (X_1, X_2, ..., X_n)$为n维随机变量，Σ为X的协方差矩阵，则Σ性质如下：

- 对称半正定
- $\sigma_{ii} = D(X_i)$
- $\sigma_{ij}^2 \le \sigma_{ii}\sigma_{jj}$

那么，我们可以给出n维正态分布的向量表示：对于$(X_1, X_2, ..., X_n) \sim N(\mu, \Sigma)$，其中$\mu = (\mu_1, \mu_2, ..., \mu_n)$是$n$维期望向量，记$X = (X_1, X_2, ..., X_n)$，则

$$f(X) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma|}} e^{-\frac{1}{2}(X-\mu)^T \Sigma^{-1}(X-\mu)}$$

若考虑到$Y = AX$，则$\Sigma(Y) = A\Sigma(X)A^T$

五・概率极限理论

5.1 随机变量序列的收敛性

定义 5.1.1（随机变量序列的收敛性）设$X, X_1, X_2, ...$为随机变量序列，X_n收敛于X的方式有：

- 概率为 1（几乎处处）的收敛：$P(\lim_{n\to\infty} X_n = X) = 1$
- 以概率收敛：$\forall \varepsilon > 0, \lim_{n\to\infty} P(|X_n - X| \ge \varepsilon) = 0$
- 以分布收敛：$\forall x \in R, \lim_{n\to\infty} F_{X_n}(x) = F_X(x)$，其中$F_{X_n}(x), F_X(x)$分别为$X_n, X$的分布函数。

定理 5.1.1（随机变量序列的收敛关系）设$X, X_1, X_2, ...$为随机变量序列，则X_n以概率为 1 收敛于X $\Rightarrow$ X_n以概率收敛于X $\Rightarrow$ X_n以分布收敛于X

5.2 大数定律

定理 5.2.1（重要不等式）设非负随机变量X的期望$E(X)$存在，则对于任意实数$\varepsilon > 0$，$P(X \ge \varepsilon) \le \frac{E(X)}{\varepsilon}$

推论 5.2.1（马尔可夫不等式） 设随机变量X的期望$E(X)$存在，则对于任意实数$\varepsilon > 0$，$P(|X| \geq \varepsilon) \leq \frac{E(|X|^k)}{\varepsilon^k}$

推论 5.2.2（切比雪夫不等式） 设随机变量X的期望$E(X)$和方差$D(X)$存在，则对于任意实数$\varepsilon > 0$，$P(|X - E(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}$

定义 5.2.1（大数定律） 设$X_1, X_2, \ldots$为随机变量序列，每个$E(X_k)$都存在，记$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$，若$\forall \varepsilon > 0, \lim_{n\to\infty} P\left(|\overline{X_n} - E\left(\overline{X_n}\right)| \geq \varepsilon\right) = 0$，则称序列$X_1, X_2, \ldots$满足大数定律。

定理 5.2.2（伯努利大数定律） 设n_A是n次独立重复实验中事件A发生的次数，$p = P(A)$，则对于任意实数$\varepsilon > 0$，$\lim_{n\to\infty} P\left(|\frac{n_A}{n} - p| \geq \varepsilon\right) = 0$

定理 5.2.3（切比雪夫大数定律） 设$X_1, X_2, \ldots$为随机变量序列，满足：1.相互独立；2.具有相同的期望和方差，则对于任意实数$\varepsilon > 0$，$\lim_{n\to\infty} P\left(|\overline{X_n} - E\left(\overline{X_n}\right)| \geq \varepsilon\right) = 0$。当然可以拓展，当相互独立但不具备相同的期望和方差时，可设$E(X_k) = \mu_k, D(X_k) = \sigma_k^2 \leq \sigma^2$，仍然有$\lim_{n\to\infty} P\left(|\overline{X_n} - E\left(\overline{X_n}\right)| \geq \varepsilon\right) = 0$。若去掉独立，则代之以$\lim_{n\to\infty} \frac{1}{n^2} D\left(\sum_{k=1}^{n} X^k\right) = 0$。

定理 5.2.4（辛钦大数定律） 设$X_1, X_2, \ldots$为随机变量序列，满足：1.相互独立；2.具有相同的分布；3.期望存在，则对于任意实数$\varepsilon > 0$，$\lim_{n\to\infty} P\left(|\overline{X_n} - E\left(\overline{X_n}\right)| \geq \varepsilon\right) = 0$

定义 5.2.2（强大数定律） 设$X_1, X_2, \ldots$为随机变量序列，每个$E(X_k)$都存在，记$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$，若$P\left(\lim_{n\to\infty} \overline{X_n} = E\left(\overline{X_n}\right)\right) = 1$，则称序列$X_1, X_2, \ldots$满足强大数定律。

定理 5.2.5（柯尔莫哥洛夫强大数定律） 设$X_1, X_2, \ldots$为随机变量序列，

- 满足：1.相互独立；2.具有相同的分布；3.期望存在，则$P\left(\lim_{n\to\infty} \overline{X_n} = E\left(\overline{X_n}\right)\right) = 1$
- 满足：1.相互独立；2.期望方差存在；3.满足柯尔莫哥洛夫条件$\sum_{n=1}^{\infty} \frac{D(X_n)}{n^2}$收敛，则$P\left(\lim_{n\to\infty} \overline{X_n} = E\left(\overline{X_n}\right)\right) = 1$

定理 5.2.6（波雷尔强大数定律） 设n_A是n次独立重复实验中事件A发生的次数，$p = P(A)$，则$P\left(\lim_{n\to\infty} \frac{n_A}{n} = p\right) = 1$

5.3 中心极限定理

定理 5.3.1（中心极限定理） 设$X_1, X_2, \ldots$为随机变量序列，若满足：期望和方差存在，且对于任意实数x，$\lim_{n\to\infty} P(\frac{\sum_{k=1}^{n} X_k - E(\sum_{k=1}^{n} X_k)}{\sqrt{D(\sum_{k=1}^{n} X_k)}} = \Phi(x)$，其中$\Phi(x)$为标准正态分布函数。则称序列$X_1, X_2, \ldots$满足中心极限定理。

定理 5.3.2（独立同分布中心极限定理） 设$X_1, X_2, \ldots$为随机变量序列，满足：1.相互独立；2.具有相同的分布；3.期望和方差存在，则对于任意实数x，$\lim_{n\to\infty} P(\frac{\sum_{k=1}^{n} X_k - nE(X_1)}{\sqrt{nD(X_1)}} = \Phi(x)$，其中$\Phi(x)$为标准正态分布函数。

定理 5.3.3（德莫佛-拉普拉斯中心极限定理） $X \sim B(n,p)$的特殊情况，近似满足$B(n,p) \sim N(np, np(1-p))$

【例】对选民进行民意调查，以预测赞成某候选人的选民比率预测比率的绝对误差不超过4.5%的概率不小于多少人？

解：设X为赞成某候选人的选民比率，$X\sim B(n,p)$，则$E(X)=p,D(X)=p(1-p)$，题目要求$P\left(\left|\frac{X}{n}-p\right|\le 0.045\right)\ge 0.95$，设$Y=\frac{X-np}{\sqrt{np(1-p)}}$，则$Y\sim N(0,1)$，而题目中$P\left(\left|\frac{X-np}{\sqrt{np(1-p)}}\right|\le\frac{0.045n}{\sqrt{np(1-p)}}\right)\ge 0.95$，则$2\Phi\left(\frac{0.045n}{\sqrt{np(1-p)}}\right)-1\ge 0.95$，故$\Phi\left(\frac{0.045n}{\sqrt{np(1-p)}}\right)\ge 0.975$，查表知$\Phi(1.96)=0.975$，因此$n\ge p(1-p)\left(\frac{1.96}{0.045}\right)^2$，故只需满足$n\ge\frac{1}{4}\left(\frac{1.96}{0.045}\right)^2=474.27$即可，故只需调查 475 人即可。

六・期中复习

6.1 知识回顾

【题型 1】考察经典分布模型的基本特征，如概率密度、期望、方差等。

- 【知识回顾】二项分布、几何分布、超几何分布、帕斯卡分布、泊松分布、均匀分布、指数分布、正态分布、伽马分布等。

1.一维

	$f_X(x)$	$E(X)$	$D(X)$
$X\sim B(n,p)$	$C_n^k p^k(1-p)^{n-k}$	np	$np(1-p)$
$X\sim G(p)$	$(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$X\sim H(N,M,n)$	$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$	$\frac{nM}{N}$	$n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$
$X\sim P(r,p)$	$C_{k-1}^{r-1}p^r(1-p)^{k-r}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
$X\sim P(\lambda)$	$e^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ
$X\sim U(a,b)$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$X\sim E(\lambda)$	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$X\sim N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2
$X\sim\Gamma(\alpha,\beta)$	$\beta^\alpha\Gamma(\alpha)^{-1}x^{\alpha-1}e^{-\beta x}$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$

2.二维/多维

	$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$	$(X,Y)\sim U(G)$
$f(X,Y)$	$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2}(1-\rho^2)^{-1}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}$	$\begin{cases}\frac{1}{\lvert G\rvert} & (x,y)\in G\\ 0 & otherwise\end{cases}$
$E(X)$	μ_1	
$D(X)$	σ_1^2	
$E(Y)$	μ_2	
$D(Y)$	σ_2^2	
$Cov(X,Y)$	$\rho\sigma_1\sigma_2$	
ρ_{XY}	ρ	
$f_X(x)$	$\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}$	
$f_Y(y)$	$\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2}$	

- 【经典例题】

【题型 2】经典概率实际问题的建模与求解。

- 【经典例题】

【题型 3】随机变量函数的分布

- 【知识回顾】

1.一维

- (1)已知随机变量X的分布，求$Y=g(X)$的分布；
- (2)已知随机变量X的分布，求$Y=aX+b$的分布；
- (3)已知随机变量X,Y的分布，求$Z=X+Y$的分布；
- (4)已知随机变量X,Y的分布，求$Z=XY$的分布；
- (5)已知随机变量X,Y的分布，求$Z=\frac{X}{Y}$的分布。

解:

- (1)

分布函数法：设X的分布函数为$F_X(x)$，则$Y=g(X)$的分布函数为$F_Y(y)=P(Y\leq y)=P(g(X)\leq y)=P(X\leq g^{-1}(y))=F_X(g^{-1}(y))$，其中$g^{-1}(y)$为$g(x)=y$的解。

公式法（要求反函数存在且可导）：$f_Y(y)=f_X(g^{-1}(y))|\frac{d}{dy}g^{-1}(y)|$

- (2)直接分布函数法求解

- (3)分布函数法：设X,Y的联合分布函数为$F_{X,Y}(x,y)$，则$Z=X+Y$的分布函数为$F_Z(z)=P(Z\leq z)=P(X+Y\leq z)=P((X,Y)\in A)$，其中$A=\{(x,y)|x+y\leq z\}$，因此

$$F_Z(z)=\iint_{(x,y)\in A} f(x,y)dxdy=\int_{-\infty}^{\infty} dx\int_{-\infty}^{z-x} f(x,y)dy=\int_{-\infty}^{z} du\int_{-\infty}^{+\infty} f(x,u-x)dx$$

。则

$$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx=\int_{-\infty}^{+\infty} f(z-y,y)dy$$

若X,Y独立，就有卷积公式

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy$$

- (4)有了(5)后，直接代$X,\frac{1}{Y}$即可
- (5)在由二维情况求出联合概率密度分布后，有$Z=\frac{X}{Y}$，令$\begin{cases}z=\frac{x}{y}\\u=y\end{cases}$，则$\begin{cases}x=zu\\y=u\end{cases}$，则$|J|=\det\left(\begin{vmatrix}z_u & z_u\\s_u & s_u\end{vmatrix}\right)=|u|$则

$$f_Z(z)=\int_{-\infty}^{+\infty} f_{ZU}(z,u)du=\int_{-\infty}^{+\infty} f(zu,u)|u|du$$

2.二维/多维

- (1)已知随机变量(X,Y)的分布，求$\begin{cases}Z=g(X,Y)\\W=h(X,Y)\end{cases}$的分布；

解:

- 分布函数法：设(X,Y)的联合分布函数为$F_{X,Y}(x,y)$，则$\begin{cases}Z=g(X,Y)\\W=h(X,Y)\end{cases}$的分布函数为$F_{Z,W}(z,w)=P(Z\le z,W\le w)=P(g(X,Y)\le z,h(X,Y)\le w)=P((X,Y)\in A)$，其中$A=\{(x,y)|g(x,y)\le z,h(x,y)\le w\}$，因此$F_{Z,W}(z,w)=\iint_{(x,y)\in A} f_{X,Y}(x,y)dxdy$。
- 公式法（要求反函数存在且连续可导）：设反函数$\begin{cases}X=u(Z,W)\\Y=v(Z,W)\end{cases}$，则$\begin{cases}Z=g(X,Y)\\W=h(X,Y)\end{cases}$的概率密度为$f_{Z,W}(z,w)=f_{X,Y}(u(z,w),v(z,w))|\frac{\partial(u,v)}{\partial(z,w)}|$，其中$\frac{\partial(u,v)}{\partial(z,w)}=\begin{vmatrix}u_z & u_w\\v_z & v_w\end{vmatrix}$为雅可比行列式。

- 【经典例题】

【题型 4】随机变量的数字特征的相互关系

- 【知识回顾】

1.联合分布与边缘分布的互相转化

- 二维分布函数与边缘分布函数：

$$F_X(x)=F(x,+\infty),F_Y(y)=F(+\infty,y)$$

- 离散型：若$p_{ij}=P\big(X=x_i,Y=y_j\big)$，则

$$p_{i\cdot}=\sum_j p_{ij},p_{\cdot j}=\sum_i p_{ij}$$

且

$$F(x,y)=\sum_{x_i\le x}\sum_{y_j\le y}p_{ij}$$

反过来

$p_{ij}=F\left(x_i,y_j\right)-F\left(x_i-0,y_j\right)-F\left(x_i,y_j-0\right)+F\left(x_i-0,y_j-0\right)$

- 连续型：若$f(x,y)$为联合密度，则

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy,f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx$$

且

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dvdu,f(x,y)=\frac{\partial^2F}{\partial x\partial y}$$

（连续点处）

2. 期望，方差，协方差等相关的等式、不等式

- 期望与线性：

$$E(aX+bY+c)=aE(X)+bE(Y)+c$$

- 二维函数期望：

$$E(g(X,Y))=\sum_i\sum_j g\left(x_i,y_j\right)p_{ij}$$

或

$$E(g(X,Y))=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dydx$$

- 方差与协方差基本式：

$$D(X)=E(X^2)-[E(X)]^2,D(aX+b)=a^2D(X)$$

$Cov(X,Y)=E(XY)-E(X)E(Y),Cov(a_1X+b_1,a_2Y+b_2)=a_1a_2Cov(X,Y)$

- 和差方差公式：

$$D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$$

若X,Y独立，则

$$Cov(X,Y)=0$$

，故

$$D(X+Y)=D(X)+D(Y)$$

- 常用不等式：

$$[E(XY)]^2\le E(X^2)E(Y^2)$$

（施瓦尔茨）

$$Cov(X,Y)^2 \le D(X)D(Y)$$

$$P(|X-E(X)| \ge \varepsilon) \le \frac{D(X)}{\varepsilon^2}$$

（切比雪夫）

- 【经典例题】

【题型 5】不同随机变量之间的关系（独立，条件）

- 【知识回顾】

1.独立

- 事件独立：$P(AB)=P(A)P(B)$，等价于（当$P(B)>0$）$P(A|B)=P(A)$。
- 二维随机变量独立的等价刻画：

$$F(x,y)=F_X(x)F_Y(y)$$

（分布函数）

$$p_{ij}=p_{i\cdot}p_{\cdot j}$$

（离散型）

$$f(x,y)=f_X(x)f_Y(y)$$

（连续点处）

- 判定常用结论： 若$f(x,y)=r(x)g(y)$（且可积、非负），则X,Y独立； 对二维正态分布，X,Y独立$\Leftrightarrow$ $\rho=0$。
- 独立的推论：

X,Y独立$\Rightarrow u(X),v(Y)$独立；

X,Y独立$\Rightarrow E(XY)=E(X)E(Y),Cov(X,Y)=0,\rho_{XY}=0$（反之一般不成立）。

2.条件

- 离散型条件分布：

$$P\big(Y=y_j|X=x_i\big)=\frac{p_{ij}}{p_{i\cdot}},P\big(X=x_i|Y=y_j\big)=\frac{p_{ij}}{p_{\cdot j}}$$

- 连续型条件密度：

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)},f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$$

- 乘法公式： 事件形式：

$$P(AB)=P(A)P(B|A)=P(B)P(A|B)$$

密度形式：

$$f(x,y)=f_X(x)f_{Y|X}(y|x)=f_Y(y)f_{X|Y}(x|y)$$

- 全概率与贝叶斯： 事件全概率：

$$P(A)=\sum_i P(B_i)P(A|B_i)$$

（B_i为A的一个划分） 事件贝叶斯：

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum_j P(B_j)P(A|B_j)}$$

密度全概率：

$$f_X(x)=\int_{-\infty}^{+\infty} f_Y(y)f_{X|Y}(x|y)dy$$

密度贝叶斯：

$$f_{X|Y}(x|y)=\frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}$$

- 独立与条件的关系： X,Y独立$\Leftrightarrow f_{Y|X}(y|x)=f_Y(y)$（或离散时$P(Y=y_j|X=x_i)=P(Y=y_j)$）。

【题型 6】概率极限理论

- 【知识回顾】

1.随机变量序列的收敛性

- 三种常见收敛： 几乎处处收敛：

$$P\left(\lim_{n\to\infty} X_n=X\right)=1$$

依概率收敛：

$$\forall\varepsilon>0,\ \lim_{n\to\infty} P(|X_n-X|\geq\varepsilon)=0$$

依分布收敛：

$$\lim_{n\to\infty} F_{X_n}(x)=F_X(x)$$

（在F_X连续点）

- 蕴含关系：

$$X_n\to X$$

(几乎处处)$\Rightarrow X_n\to X$(依概率)$\Rightarrow X_n\to X$(依分布)

- 常用判据（切比雪夫）：

$$P(|X_n-X|\geq\varepsilon)\leq\frac{D(X_n-X)}{\varepsilon^2}$$

，若$D(X_n - X) \to 0$，则$X_n \to X$（依概率）。

2.大数定律

- 样本均值：

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

- 伯努利大数定律：

$$\frac{n_A}{n} \to p$$

（依概率）

- 切比雪夫大数定律（常用版）：若$X_1, X_2, ...$相互独立，且同均值同方差，则

$$\overline{X_n} - E\left(\overline{X_n}\right) \to 0$$

（依概率）

- 辛钦大数定律：若$X_1, X_2, ...$独立同分布且$E(X_1)$存在，则

$$\overline{X_n} \to E(X_1)$$

（依概率）

- 强大数定律（记忆版）：若$X_1, X_2, ...$独立同分布且$E(X_1)$存在，则

$$P\left(\lim_{n\to\infty} \overline{X_n} = E(X_1)\right) = 1$$

3.中心极限定理与正态近似

- 独立同分布中心极限定理：若$E(X_i) = \mu, D(X_i) = \sigma^2$，则

$$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \to N(0,1)$$

等价地

$$\frac{\overline{X_n} - \mu}{\frac{\sigma}{\sqrt{n}}} \to N(0,1)$$

- 二项分布正态近似（德莫佛-拉普拉斯）： $X \sim B(n,p), n$大且p不过分接近$0,1$时，

$$X \approx N(np, np(1-p))$$

- 连续性修正（高频）：

$$P(a \le X \le b) \approx \Phi\left(\frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - 0.5 - np}{\sqrt{np(1-p)}}\right)$$

- 做题流程：标准化→ Φ表→(离散型时)连续性修正→必要时反解样本量n。

6.2 lds 老师班 23 往年题（部分）

【T1】在圆周上随机选取三点，这三点构成的三角形是钝角三角形的概率是？

解：将圆周长视为 1，三个随机点将圆周分为三段弧长x, y, z，且$x + y + z = 1$。这等价于在$\{(x, y)|x > 0, y > 0, x + y < 1\}$上随机选取一个点$(x, y)$。三角形是钝角三角形的充要条件是：$\max\{x, y, z\} > \frac{1}{2}$，即$x > \frac{1}{2}$或$y > \frac{1}{2}$或$x + y < \frac{1}{2}$，因此所求概率为$\frac{3}{4}$。

【T2】设X, Y都服从正态分布，且它们不相关则它们一定独立吗？

解：未必，(X, Y)不服从二维正态分布时，不相关不一定独立。

注：容易与X, Y相互独立$\Leftrightarrow$ $\rho_{XY} = 0$对于(X, Y)二维正态分布的情况成立混淆。

【T3】设$X \sim N(1, 1), Y \sim N(2, 1)$，且$\rho_{XY} = -\frac{1}{2}$，求$E[(X + Y)^2]$。

解：$E[(X + Y)^2] = E(X^2) + E(Y^2) + 2E(XY)$，而已知$E(X) = 1, E(Y) = 2, D(X) = D(Y) = 1, \rho_{XY} = -\frac{1}{2} = \frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}$，则$E(XY) = \rho_{XY}\sqrt{D(X)D(Y)}) + E(X)E(Y) = \frac{3}{2}$，又$E(X^2) = D(X) + [E(X)]^2 = 2, E(Y^2) = D(Y) + [E(Y)]^2 = 5$，因此$E[(X + Y)^2] = 2 + 5 + 2\left(\frac{3}{2}\right) = 10$

【T4】设A是已知的$n \times n$矩阵，$X = (X_1, X_2, ..., X_n)^T$是$n$维随机变量，且服从$n$维标准正态分布，求$E(X^T AX)$。

解：n维标准正态分布意味着$\mu = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \Sigma = I_n$，也就意味着$E(X_i X_j) = 0$当$i \neq j, E(X_i^2) = 1$，因此$E(X^T AX) = E\left(\sum_{i=1}^n \sum_{j=1}^n a_{ij} X_i X_j\right) = \sum_{i=1}^n a_{ii} E(X_i^2) = \mathrm{tr}(A)$

【T5】设$E(X) = E(Y) = 1, \Sigma(X, Y) = \begin{pmatrix} 9 & 1 \\ 1 & 16 \end{pmatrix}$，根据切比雪夫不等式，求$P(|X - Y| \geq 25)$的上界。

解：$P(|X - Y| \geq 25) = P((X - Y)^2 \geq 25^2) \leq \frac{E((X-Y)^2)}{25^2} = \frac{23}{625}$

【T6】设$X_1, X_2, ...$为独立同分布的随机变量，且都服从伯努利分布$B(n, p)$，则几乎处处有$\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^k X_i^2 = ?$

解：$\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^k X_i^2 = E(X_i^2) = D(X) + [E(X)]^2 = np(1 - p) + n^2 p^2 = np[1 + (n - 1)p]$

【T7】某饮品厂商推出了“56 个民族特别版盲盒”。每开一个盲盒，会随机等概率获得 56 个款式之一的罐装饮料。要集齐所有 56 个款式，需要开盲盒的数目期望最接近下面哪个数？（250，350，450，550）

解：记X_i为集齐$i - 1$个不同款式之后，到集齐i个不同款式需要开盲盒的数目，则X_i服从几何分布$G\left(\frac{56-(i-1)}{56}\right)$，因此$E(X_i) = \frac{56}{56-(i-1)}$，因此集齐所有 56 个款式需要开盲盒的数目的期望为$E\left(\sum_{i=1}^{56} X_i\right) = \sum_{i=1}^{56} E(X_i) = 56 \sum_{i=1}^{56} \frac{1}{i}$而$\sum_{i=1}^{n} \frac{1}{i} \in (\ln(n + 1), \ln n + \ln 2)$故$E\left(\sum_{i=1}^{56} X_i\right) = 56 \sum_{i=1}^{56} \frac{1}{i} \approx 56(\ln 56 + \ln 2) \approx 56(3.99 + 0.69) \approx 264.48$。

【T8】某地区患有某种疾病的人占0.05%（注意这里有百分号），患者对某种检测结果呈阳性的概率为99%，正常人对这种检测呈阳性的概率为0.5%，求检测结果呈阳性的人得这种病的概率（结果只需要精确到形如 n%，其中 n 是非负整数）。

解：设A为患病事件，B为检测结果呈阳性事件，则题目要求$P(A|B)$，根据贝叶斯公式，$P(A|B)=\frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)P(A^c)}=\frac{0.99*0.0005}{0.99*0.0005+0.005*0.9995}\approx 9.01\%$

【T9】设随机变量X的概率密度$f_X(x)=\begin{cases}x & 0\le x\le 1\\ 2-x & 1<x\le 2\\ 0 & \text{otherwise}\end{cases}$

- (1)求a使得$P(X\le a)=3P(X>a)$
- (2)求$Y=X^2$的概率密度函数$f_Y(y)$

解：

- (1)画出图像，$a=2-\frac{\sqrt{2}}{2}$
- (2)使用分布函数法，$F_Y(y)=P(Y\le y)=P(X^2\le y)=P(-\sqrt{y}\le X\le\sqrt{y})$当$0\le y\le 1$时，$F_Y(y)=\int_0^{\sqrt{y}}f_X(x)dx=\int_0^{\sqrt{y}}xdx=\frac{y}{2}$当$1<y\le 4$时，$F_Y(y)=\int_0^1 f_X(x)dx+\int_1^{\sqrt{y}}f_X(x)dx=\frac{1}{2}+\int_1^{\sqrt{y}}(2-x)dx=\frac{3}{2}-\sqrt{y}+\frac{y}{2}$因此$f_Y(y)=\begin{cases}\frac{1}{2} & 0<y\le 1\\ \frac{1}{\sqrt{y}}-\frac{1}{2} & 1<y\le 4\\ 0 & \text{otherwise}\end{cases}$

【T10】设昆虫产卵个数X服从参数为λ的泊松分布，又设一个虫卵能孵化为昆虫的概率等于p，已知虫的孵化是互相独立的，记下一代昆虫的个数为Y。

- (1)求X,Y的联合分布以及Y的边缘分布
- (2)证明Y和$X-Y$相互独立，判断Y和$X-Y$是否不相关

解：

- (1)由题得$P(X=n)=\frac{\lambda^n}{n!}e^{-\lambda},P(Y=k|X=n)=C_n^k p^k(1-p)^{n-k}$，则$P(X=n,Y=k)=P(Y=k|X=n)P(X=n)=C_n^k p^k(1-p)^{n-k}\frac{\lambda^n}{n!}e^{-\lambda}$，因此$Y$的边缘分布为$P(Y=k)=\sum_{n=k}^{\infty}P(X=n,Y=k)=\sum_{n=k}^{\infty}C_n^k p^k(1-p)^{n-k}\frac{\lambda^n}{n!}e^{-\lambda}=\frac{(p\lambda)^k}{k!}e^{-p\lambda}$
- (2)$P(Y=k,X-Y=j)=P(Y=k,X=k+j)=P(X=k+j,Y=k)=C_{k+j}^k p^k(1-p)^j\frac{\lambda^{k+j}}{(k+j)!}e^{-\lambda}$，而$P(Y=k)P(X-Y=j)=\frac{(p\lambda)^k}{k!}e^{-p\lambda}*(1-p)^j\frac{\lambda^j}{j!}e^{-\lambda(1-p)}$，通过计算可以验证两者相等，因此$Y$和$X-Y$相互独立。由于它们相互独立，因此它们不相关。

【T11】抢红包问题的背景简述如下：$X_1\sim U(0,\frac{2s}{n})$，且对于$k=2\sim n-1$，$X_k\mid(X_1,...,X_{k-1})\sim U\left(0,\frac{2\sum_{i=1}^{k-1}x_i}{n-k+1}\right),X_n=s-\sum_{i=1}^{n-1}X_i$。

- (1)$n=2$时，是否有$X_1=X_2$？
- (2)$n\ge 3$时，求X_2的概率密度。
- (3)$n\ge 3$时，求$E(X_1),E(X_2)$大小关系。

解：

- (1)只能得出$X_1\sim U(0,s),X_1+X_2=s$，而$X_1=X_2=\frac{s}{2}$为一定值，因此$X_1=X_2$的概率为0。它们同分布不相等。
- (2)由题得

$$X_1 \sim U\left(0, \frac{2s}{n}\right)$$

$$X_2|X_1 \sim U\left(0, \frac{2X_1}{n-1}\right)$$

则

$$f_{X_1}(x_1) = \frac{1}{2\frac{s}{n}} = \frac{n}{2s}$$

$$f_{X_2|X_1}(x_2|x_1) = \frac{1}{2\frac{x_1}{n-1}} = \frac{n-1}{2x_1}$$

则

$$f(x_1, x_2) = f_{X_1}(x_1) f_{X_2|X_1}(x_2|x_1) = \frac{n(n-1)}{4sx_1}, 0 < x_1 < \frac{2s}{n}, 0 < x_2 < \frac{2x_1}{n-1}$$

则

$$f_{X_2}(x_2) = \int_{\frac{(n-1)x_2}{2}}^{\frac{2s}{n}} f(x_1, x_2) dx_1 = \frac{n(n-1)}{4s} \ln\left(\frac{4s}{n(n-1)x}\right), 0 < x < \frac{4s}{n(n-1)}$$

6.3 wrs 老师班

6.4 copilot 押题

说明：本套押题覆盖一到五章核心内容，难度高于常规期中卷；默认可使用标准正态分布表。

【T1】 设X_1, X_2, X_3独立同分布且$X_i \sim U(0,1)$，记$R = \max(X_1, X_2, X_3) - \min(X_1, X_2, X_3)$。

- (1)求R的分布函数与密度函数。
- (2)求$P\left(R < \frac{1}{3}\right)$。
- (3)求$E(R), D(R)$。

解：

- (1)对$0 < r < 1$，有$F_{R(r)} = P(R \le r) = 3r^2 - 2r^3$，故

$$F_{R(r)} = \begin{cases} 0 & r \le 0 \\ 3r^2 - 2r^3 & 0 < r < 1 \\ 1 & r \ge 1 \end{cases}$$

从而

$$f_{R(r)} = F_{R'}(r) = 6r(1-r), 0 < r < 1$$

。

- (2)

$$P\left(R < \frac{1}{3}\right) = F_{R(\frac{1}{3})} = 3 * \left(\frac{1}{9}\right) - 2 * \left(\frac{1}{27}\right) = \frac{7}{27}$$

  。

- (3)

$$E(R) = \int_0^1 r f_{R(r)} dr = \frac{1}{2}, E(R^2) = \int_0^1 r^2 f_{R(r)} dr = \frac{3}{10}, D(R) = E(R^2) - E(R)^2 = \frac{1}{20}$$

  。

【T2】某疾病患病率为0.2%。先做检测A，再对A阳性者做检测B。已知

- $P(A+|D) = 0.98, P(A+|D^c) = 0.03$
- $P(B+|D) = 0.99, P(B+|D^c) = 0.005$

并且在给定D或D^c条件下，两次检测结果相互独立。

- (1)求$P(D|A+, B+)$。
- (2)求$P(D|A+, B-)$。

解：记$P(D) = 0.002, P(D^c) = 0.998$。

- (1)

$$P(A+, B+|D) = 0.98 * 0.99 = 0.9702, P(A+, B+|D^c) = 0.03 * 0.005 = 0.00015$$

  。

因此

$$P(D|A+, B+) = \frac{0.9702 * 0.002}{0.9702 * 0.002 + 0.00015 * 0.998} \approx 0.9284$$

。

- (2)

$$P(A+, B-|D) = 0.98 * 0.01 = 0.0098, P(A+, B-|D^c) = 0.03 * 0.995 = 0.02985$$

  。

故

$$P(D|A+, B-) = \frac{0.0098 * 0.002}{0.0098 * 0.002 + 0.02985 * 0.998} \approx 6.58 \times 10^{-4} \approx 0.0658\%$$

。

【T3】设$X \sim P(\lambda)$。在给定X条件下，每个个体独立地以概率p被系统识别，识别总数记U；另有独立背景噪声$Z \sim P(\mu)$。观测量$Y = U + Z$。

- (1)求Y的分布。
- (2)求$U|Y = n$的条件分布。
- (3)求$E(X|Y = n)$。

解：

- (1)泊松稀释得$U \sim P(\lambda p)$，且$V := X - U \sim P(\lambda(1-p))$，并有$U, V$独立。又$Z \sim P(\mu)$独立，故

$$Y = U + Z \sim P(\lambda p + \mu)$$

。

- (2)由“独立泊松和给定总和的条件分布”为二项分布，得

$$U|Y = n \sim B\left(n, \frac{\lambda p}{\lambda p + \mu}\right)$$

。

- (3)

$$E(X|Y = n) = E(U + V|Y = n) = E(U|Y = n) + E(V) = n\frac{\lambda p}{\lambda p + \mu} + \lambda(1-p)$$

。

【T4】设随机变量X的密度为

$$f_{X(x)} = 2x, 0 < x < 1$$

（其余为 0），令$Y = \ln \frac{X}{1-X}$。

- (1)求Y的密度函数。
- (2)求$E(Y)$。
- (3)求$P(Y > 0)$。

解：由

$$y = \ln \frac{x}{1-x}$$

得

$$x = \frac{e^y}{1+e^y}, y \in R$$

，且

$$\frac{dx}{dy} = \frac{e^y}{(1+e^y)^2}$$

。

- (1)

$$f_{Y(y)}=f_{X(x(y))}\left|\frac{dx}{dy}\right|=2\frac{e^y}{1+e^y}\frac{e^y}{(1+e^y)^2}=\frac{2e^{2y}}{(1+e^y)^3},y\in R$$

。
- (2)

$$E(Y)=E(\ln X)-E(\ln(1-X))$$

，

其中

$$E(\ln X)=\int_0^1 2x\ln x dx=-\frac{1}{2}$$

，

$$E(\ln(1-X))=\int_0^1 2x\ln(1-x)dx=-\frac{3}{2}$$

。 故$E(Y)=1$。
- (3)$Y>0\Leftrightarrow X>\frac{1}{2}$，故

$$P(Y>0)=\int_{\frac{1}{2}}^1 2xdx=\frac{3}{4}$$

。

【T5】设二维随机变量(X,Y)的联合密度为

$$f(x,y)=3(x+y),\quad x>0,y>0,x+y<1$$

其余为0。
- (1)求边缘密度f_X,f_Y，并判断X,Y是否独立。
- (2)求$P\left(Y<\frac{X}{2}\right)$。
- (3)求$E(Y|X=x)$。
- (4)求$Cov(X,Y)$。

解:
- (1)

$$f_{X(x)}=\int_0^{1-x}3(x+y)dy=\frac{3}{2}(1-x^2),0<x<1$$

； 同理

$$f_{Y(y)}=\frac{3}{2}(1-y^2),0<y<1$$

。由于支持域为三角形$x+y<1$，并非直积区间，故X,Y不独立。

- (2)按y上界分段：

$$P\left(Y<\frac{X}{2}\right)=\int_0^{\frac{2}{3}}\int_0^{\frac{x}{2}}3(x+y)dydx+\int_{\frac{2}{3}}^{1}\int_0^{1-x}3(x+y)dydx=\frac{1}{3}$$

。

- (3)

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_{X(x)}}=\frac{2(x+y)}{1-x^2},0<y<1-x$$

，

所以

$$E(Y|X=x)=\int_0^{1-x}y\frac{2(x+y)}{1-x^2}dy=\frac{(1-x)(x+2)}{3(1+x)}$$

。

- (4)

$$E(X)=E(Y)=\frac{3}{8},E(XY)=\int_0^1\int_0^{1-x}3xy(x+y)dydx=\frac{1}{10}$$

，

故

$$Cov(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{10}-\left(\frac{3}{8}\right)^2=-\frac{13}{320}$$

。

【T6】设X,Y相互独立且都服从指数分布$E(\lambda)$，令

$$U=X+Y,V=\frac{X}{X+Y}$$

。

- (1)求(U,V)联合密度。
- (2)判断U,V是否独立，并给出各自分布。
- (3)求$P(X>2Y)$。

解：令

$$x=uv,y=u(1-v)$$

，则$u>0,0<v<1$，雅可比$|J|=u$。

- (1)

$$f_{UV}(u,v)=\lambda e^{-\lambda uv}\lambda e^{-\lambda u(1-v)}u=\lambda^2 ue^{-\lambda u},u>0,0<v<1$$

。

- (2)上式可分解为$[\lambda^2 ue^{-\lambda u}]*[1]$，故$U,V$独立；

其中

$$U\sim\Gamma(2,\lambda)$$

，

$$V\sim U(0,1)$$

。

- (3)

  $$X>2Y\Leftrightarrow V>\frac{2}{3}$$

  ，故

  $$P(X>2Y)=P\left(V>\frac{2}{3}\right)=\frac{1}{3}$$

  。

【T7】设(X,Y)服从二维正态分布，参数为 $\mu_1=0,\sigma_1^2=1,\mu_2=0,\sigma_2^2=4,\rho=\frac{1}{2}$。
- (1)求$Z=2X-Y$的分布。
- (2)求$X|Y=y$的条件分布。
- (3)求$P(X>0,Y>0)$。

解：

$$Cov(X,Y)=\rho\sigma_1\sigma_2=1$$

。

- (1)

  $$E(Z)=0,D(Z)=4D(X)+D(Y)-4Cov(X,Y)=4+4-4=4$$

  ，故

  $$Z\sim N(0,4)$$

  。

- (2)二维正态条件分布公式：

$$X|Y=y\sim N\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2)\right)$$

即

$$X|Y=y\sim N\left(\frac{y}{4},\frac{3}{4}\right)$$

。

- (3)二维正态象限概率公式：

$$P(X>0,Y>0)=\frac{1}{4}+\frac{1}{2\pi}\arcsin(\rho)=\frac{1}{4}+\frac{1}{2\pi}*\frac{\pi}{6}=\frac{1}{3}$$

。

【T8】设$X=(X_1,X_2,X_3)^T\sim N(0,I_3)$，

$$A=\begin{pmatrix}2&1&0\\1&2&0\\0&0&1\end{pmatrix}$$

，定义$Q=X^TAX$。

- (1)求$E(Q),D(Q)$。
- (2)写出Q可表示成若干独立$\chi^2(1)$线性组合的形式。
- (3)写出Q的矩母函数$M_{Q(t)}$。

解：

- (1)

  $E(Q)=\mathrm{tr}(A)=5$

  ；

$$D(Q)=2\,\mathrm{tr}(A^2)$$

（标准正态二次型结论）。由A特征值为$3,1,1$，故$\mathrm{tr}(A^2)=3^2+1^2+1^2=11$，所以

$$D(Q)=22$$

。

- (2)设Z_1,Z_2,Z_3独立且$Z_i\sim N(0,1)$，则

Q与随机变量$3Z_1^2+Z_2^2+Z_3^2$同分布。

- (3)

$$M_{Q(t)}=\prod_{i=1}^{3}(1-2\lambda_i t)^{-\frac{1}{2}}=(1-6t)^{-\frac{1}{2}}(1-2t)^{-1},\quad t<\frac{1}{6}$$

。

【T9】设随机变量X,Y满足$E(X)=0,E(Y)=1,D(X)=4,D(Y)=9,Cov(X,Y)=3$。

- (1)求相关系数ρ_{XY}。
- (2)求使$E[(Y-aX-b)^2]$最小的a,b。
- (3)给出该最小值。

解：

- (1)

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}} = \frac{3}{2*3} = \frac{1}{2}$$

。

- (2)最小二乘投影系数

$$a^\star = \frac{Cov(X,Y)}{D(X)} = \frac{3}{4}, b^\star = E(Y) - a^\star E(X) = 1$$

。

- (3)最小均方误差

$$D(Y) - \frac{Cov(X,Y)^2}{D(X)} = 9 - \frac{9}{4} = \frac{27}{4}$$

。

【T10】设$X_1, X_2, ...$相互独立，且$E(X_n) = 0, D(X_n) = \frac{1}{n^\alpha}$，记

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

。

- (1)求使$\overline{X_n} \to 0$（依概率）成立的α取值。
- (2)给出使$\overline{X_n} \to 0$（几乎处处）成立的一个充分条件。

解：

- (1)

$$D\left(\overline{X_n}\right) = \frac{1}{n^2}\sum_{k=1}^{n}\frac{1}{k^\alpha}$$

。 当$\alpha > -1$时，上式$\to 0$，由切比雪夫不等式得$\overline{X_n} \to 0$（依概率）。 当$\alpha = -1$时，方差趋于正常数；$\alpha \leftarrow 1$时方差发散，因此不依概率收敛到 0。 故充要条件是$\alpha > -1$。

- (2)若

$$\alpha > -1$$

，则

$$\sum_{n=1}^{\infty}\frac{D(X_n)}{n^2} = \sum_{n=1}^{\infty}\frac{1}{n^{\alpha+2}} < \infty$$

， 由柯尔莫哥洛夫强大数定律可得$\overline{X_n} \to 0$几乎处处成立。

【T11】设到达数$N \sim P(180)$。

- (1)用中心极限定理近似计算$P(160 \le N \le 210)$（使用连续性修正）。

- (2)求最小整数m，使得$P(N \le m) \approx 0.99$。

解：记$\mu = 180, \sigma = \sqrt{180} \approx 13.416$。

- (1)

$$P(160 \le N \le 210) \approx P(159.5 \le Y \le 210.5), Y \sim N(180, 180)$$

因此

$$P \approx \Phi\left(\frac{210.5 - 180}{13.416}\right) - \Phi\left(\frac{159.5 - 180}{13.416}\right) = \Phi(2.27) - \Phi(-1.53) \approx 0.9884 - 0.0634 = 0.9250$$

。

- (2)令

$$P(N \le m) \approx \Phi\left(\frac{m + 0.5 - 180}{13.416}\right) = 0.99$$

，取$z_{0.99} = 2.326$，得

$$m + 0.5 \approx 180 + 2.326 * 13.416 \approx 211.2$$

，故最小整数$m = 211$。

【T12】设$U \sim U(-1, 1)$，定义$X = U, Y = U^2$。

- (1)求$Cov(X, Y)$与ρ_{XY}。
- (2)判断X, Y是否独立并说明理由。

解：

- (1)

$$E(X) = E(U) = 0, E(Y) = E(U^2) = \frac{1}{3}, E(XY) = E(U^3) = 0$$

，故

$$Cov(X, Y) = E(XY) - E(X)E(Y) = 0$$

。又

$$D(X) = \frac{1}{3}, D(Y) = E(U^4) - E(U^2)^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$$

，故$\rho_{XY} = 0$。

- (2)不独立。取事件

$$A = \left\{|X| > \frac{1}{2}\right\}, B = \left\{Y > \frac{1}{4}\right\}$$

，由$Y = X^2$可知$A = B$，于是

$$P(A \cap B) = P(A) = \frac{1}{2}, P(A)P(B) = \frac{1}{4}$$

，二者不等，故X,Y不独立。

6.5 题型总结

说明：以下按“知识点/题型 -> 经典例题 -> 解题关键词”整理，覆盖期中常见与高频变式；建议按章节先做基础题型，再做综合题型。

A. 随机事件与概率（第一章）

【题型 01】事件运算与德摩根律

- 知识点：并、交、差、补，集合恒等变形。
- 经典例题：设$P(A)=0.6, P(B)=0.5, P(A\cap B)=0.3$，求$P\left(\overline{A}\cup\overline{B}\right)$。

解题关键词：先用德摩根$\overline{A}\cup\overline{B}=\overline{A\cap B}$，再用$1-P(A\cap B)$。

【题型 02】古典概型与排列组合计数

- 知识点：等可能、排列组合、分步计数。
- 经典例题：8 个球中 5 红 3 白，不放回取 2 个，求“恰有 1 红 1 白”的概率。

解题关键词：$\frac{\binom{5}{1}\binom{3}{1}}{\binom{8}{2}}$。

【题型 03】几何概型（面积法）

- 知识点：二维均匀落点，概率=面积比。
- 经典例题：(X,Y)在单位正方形上均匀分布，求$P\left(X+Y<\frac{1}{2}\right)$。

解题关键词：目标区域是直角三角形，面积$=\frac{1}{8}$。

【题型 04】加法公式与容斥原理

- 知识点：$P(A\cup B)$、多事件容斥。
- 经典例题：从 1 到 100 随机取整数，求能被 2 或 3 或 5 整除的概率。

解题关键词：先数个数，用容斥：$50+33+20-16-10-6+3$。

【题型 05】条件概率基础题

- 知识点：$P(A|B)=\frac{P(A\cap B)}{P(B)}$。
- 经典例题：掷两枚公平骰子，已知点数和大于 8，求“至少有一个 6”的概率。

解题关键词：在条件样本空间内重新计数。

【题型 06】全概率公式

- 知识点：完备事件组拆分。
- 经典例题：两台机器产量占比$3:2$，次品率分别为1%和2%，求随机产品为次品概率。

解题关键词：$P(D)=P(M_1)P(D|M_1)+P(M_2)P(D|M_2)$。

【题型 07】贝叶斯反推

- 知识点：先验、似然、后验。
- 经典例题：沿用题型 06，已知抽到次品，求其来自第二台机器的概率。

解题关键词：$P(M_2|D)=\frac{P(D|M_2)P(M_2)}{P(D)}$。

【题型 08】独立性判断与常见误区

- 知识点：独立≠互斥；检验$P(A\cap B)=P(A)P(B)$。
- 经典例题：设$P(A)=0.4,P(B)=0.5,P(A\cap B)=0.2$，判断$A,B$是否独立。

解题关键词：直接比较0.2与$0.4*0.5$。

B. 一维随机变量及分布（第二章）

【题型 09】伯努利/二项分布建模

- 知识点：独立重复试验，成功次数。
- 经典例题：某题正确率$p=0.7$，独立做 10 题，求恰好对 8 题概率。

解题关键词：$\binom{10}{8}0.7^8 0.3^2$。

【题型 10】几何分布与“首次成功”

- 知识点：首次成功时刻、无记忆思想。
- 经典例题：每次成功率0.2，求第 5 次首次成功概率。

解题关键词：$(1-p)^{k-1}p$。

【题型 11】帕斯卡分布与“第r次成功”

- 知识点：第r次成功出现时刻。
- 经典例题：每次成功率0.4，求第 3 次成功出现在第 7 次试验概率。

解题关键词：$\binom{6}{2}0.4^3 0.6^4$。

【题型 12】超几何分布（不放回抽样）

- 知识点：有限总体、无放回。
- 经典例题：20 件中 5 件次品，抽 4 件，求恰好 1 件次品概率。

解题关键词：$\frac{\binom{5}{1}\binom{15}{3}}{\binom{20}{4}}$。

【题型 13】泊松分布与泊松近似

- 知识点：稀有事件，$\lambda=np$。
- 经典例题：$X\sim B(1000,0.002)$，近似求$P(X=0)$。

解题关键词：$X\approx P(2)$，故$P(X=0)\approx e^{-2}$。

【题型 14】分布函数法（由F求概率）

- 知识点：$P(a < X \le b) = F(b) - F(a)$。
- 经典例题：给定分段$F(x)$，求$P(1 < X \le 3)$并判断该点是否有跳跃。

解题关键词：左极限、右连续、差分取概率。

【题型 15】由密度求参数

- 知识点：归一化$\int f = 1$。
- 经典例题：$f(x) = cx(1-x), 0 < x < 1$，求$c$。

解题关键词：$c\int_0^1 x(1-x)dx = 1$。

【题型 16】常见连续分布识别与概率计算

- 知识点：均匀、指数、正态、伽马。
- 经典例题：$X \sim E(2)$，求$P(X > 1.5)$与$E(X), D(X)$。

解题关键词：尾概率$e^{-\lambda t}$，并用$E = \frac{1}{\lambda}, D = \frac{1}{\lambda^2}$。

【题型 17】正态标准化与查表

- 知识点：$Z = \frac{X-\mu}{\sigma}$。
- 经典例题：$X \sim N(50, 9)$，求$P(47 < X < 56)$。

解题关键词：先标准化成Φ差值。

【题型 18】单调变换的密度法

- 知识点：$Y = g(X)$单调时

$$f_{Y(y)} = f_{X(x(y))}\left|d\frac{x}{d}y\right|$$

。

- 经典例题：$X \sim U(0,1)$，$Y = -\ln X$，求Y分布。

解题关键词：反函数$x = e^{-y}$，得到$Y \sim E(1)$。

【题型 19】非单调变换（分布函数法）

- 知识点：多分支逆像。
- 经典例题：$X \sim U(-1,1)$，$Y = X^2$，求$f_{Y(y)}$。

解题关键词：先算$F_{Y(y)} = P(-\sqrt{y} \le X \le \sqrt{y})$再求导。

C. 多维随机变量（第三章）

【题型 20】联合密度归一化求常数

- 知识点：二维积分归一化。
- 经典例题：$f(x,y) = c(x+y), x > 0, y > 0, x + y < 1$，求$c$。

解题关键词：在三角域上积分，结果$c=3$。

【题型 21】边缘密度
- 知识点：对另一个变量积分求边缘。
- 经典例题：对题型 20，求$f_{X(x)}, f_{Y(y)}$。

解题关键词：按支持域上限积分，注意区间$0<x<1$。

【题型 22】二维独立性判定
- 知识点：$f(x,y)=f_{X(x)}f_{Y(y)}$与支持域直积性。
- 经典例题：判断题型 20 中的X,Y是否独立。

解题关键词：支持域是三角形，不是矩形直积，通常不独立。

【题型 23】条件密度与条件期望
- 知识点：$f_{Y|X}=\frac{f}{f_X}$，再积分求$E(Y|X=x)$。
- 经典例题：对题型 20，求$f_{Y|X}(y|x)$与$E(Y|X=x)$。

解题关键词：先写条件密度，再做一维积分。

【题型 24】二维函数分布：和的分布
- 知识点：$Z=X+Y$的分布函数域积分。
- 经典例题：X,Y独立且$\sim E(\lambda)$，求$Z=X+Y$的密度。

解题关键词：卷积，结果为$\Gamma(2,\lambda)$型。

【题型 25】卷积公式通用题
- 知识点：$f_{Z(z)}=\int f_{X(x)}f_{Y(z-x)}dx$。
- 经典例题：$X\sim U(0,1), Y\sim U(0,1)$独立，求$Z=X+Y$密度。

解题关键词：分段$0<z<1$与$1\le z<2$。

【题型 26】Jacobian 变换（线性变换）
- 知识点：构造(U,V)并算雅可比。
- 经典例题：$U=X+Y, V=X-Y$，由f_{XY}求f_{UV}。

解题关键词：反解$x=\frac{u+v}{2}, y=\frac{u-v}{2}$，$|J|=\frac{1}{2}$。

【题型 27】商的分布
- 知识点：$Z=\frac{X}{Y}$常配合(Z,U)变换。
- 经典例题：设X,Y独立标准正态，求$Z=\frac{X}{Y}$的分布类型。

解题关键词：结果为标准 Cauchy 分布。

【题型 28】极值分布
- 知识点：$F_{\max}$、$F_{\min}$。

- 经典例题：X,Y独立同分布，分布函数为F，求$\max(X,Y)$分布函数。

解题关键词：$F_{\max}(z)=F(z)^2$，$F_{\min}(z)=1-(1-F(z))^2$。

【题型 29】二维正态线性组合
- 知识点：均值线性、方差含协方差项。
- 经典例题：(X,Y)二维正态已知参数，求$Z=2X-Y$分布。

解题关键词：$E(Z)=2E(X)-E(Y)$，$D(Z)=4D(X)+D(Y)-4Cov(X,Y)$。

【题型 30】二维正态条件分布与象限概率
- 知识点：$X|Y=y$仍正态；$P(X>0,Y>0)$与ρ关系。
- 经典例题：(X,Y)零均值二维正态，已知相关系数ρ，求$P(X>0,Y>0)$。

解题关键词：$\frac{1}{4}+\frac{1}{2\pi}\arcsin(\rho)$。

D. 数字特征（第四章）

【题型 31】离散/连续型期望方差直接计算
- 知识点：定义法与$D(X)=E(X^2)-E(X)^2$。
- 经典例题：X密度$f(x)=2x,0<x<1$，求$E(X),D(X)$。

解题关键词：直接积分，注意上下限。

【题型 32】随机变量函数的期望
- 知识点：LOTUS：$E(g(X))$不必先求$g(X)$分布。
- 经典例题：$X\sim U(0,1)$，求$E(X^3-\ln X)$。

解题关键词：拆成两个积分。

【题型 33】协方差与相关系数
- 知识点：$Cov=E(XY)-E(X)E(Y)$，$\rho=Co\frac{v}{\sigma_X\sigma_Y}$。
- 经典例题：已知$E(X),E(Y),D(X),D(Y),E(XY)$，求$Cov,\rho$。

解题关键词：先算Cov再标准化。

【题型 34】方差性质综合
- 知识点：$D(aX+b)=a^2D(X)$，独立时方差可加。
- 经典例题：已知$D(X)=2,D(Y)=3,Cov(X,Y)=-1$，求$D(2X-Y)$。

解题关键词：$D(2X-Y)=4D(X)+D(Y)-4Cov(X,Y)$。

【题型 35】协方差矩阵与线性变换
- 知识点：$\Sigma(AX)=A\Sigma(X)A^T$。
- 经典例题：给定$\Sigma(X)$与矩阵A，求$Y=AX$的协方差矩阵。

解题关键词：按矩阵乘法顺序计算。

【题型 36】二次型X^TAX

- 知识点：标准正态下$E(X^TAX) = \mathrm{tr}(A)$。
- 经典例题：$X \sim N(0, I_n)$，求$E(X^TAX)$；进阶求$D(X^TAX)$。

解题关键词：期望用迹；方差常用$2\,\mathrm{tr}(A^2)$（标准正态情形）。

E. 概率不等式与极限定理（第五章）

【题型 37】马尔可夫/切比雪夫不等式

- 知识点：上界估计题。
- 经典例题：已知$E(X) = 0, D(X) = 9$，求$P(|X| \geq 6)$的上界。

解题关键词：切比雪夫：$\leq \frac{9}{36} = \frac{1}{4}$。

【题型 38】收敛方式关系判定

- 知识点：几乎处处收敛⇒依概率收敛⇒分布收敛。
- 经典例题：给出一列随机变量，判断哪种收敛成立并说明蕴含关系。

解题关键词：先判最强收敛，再向下推。

【题型 39】大数定律应用

- 知识点：样本均值稳定到期望。
- 经典例题：独立同分布X_i满足$E(X_i) = \mu$，求$\frac{1}{n}\sum_{i=1}^{n} X_i$的极限（依概率或几乎处处）。

解题关键词：辛钦大数定律/强大数定律。

【题型 40】柯尔莫哥洛夫条件型强大数律

- 知识点：独立但不同分布时的充分条件。
- 经典例题：$D(X_n) = \frac{1}{n^\alpha}$，判断$\frac{1}{n}\sum_{k=1}^{n} X_k$是否几乎处处收敛到 0。

解题关键词：检查$\sum \frac{D(X_n)}{n^2}$是否收敛。

【题型 41】中心极限定理近似计算

- 知识点：和/均值的正态近似，离散型连续性修正。
- 经典例题：$N \sim P(180)$，近似求$P(160 \leq N \leq 210)$。

解题关键词：$N \approx N(180, 180)$并用159.5与210.5修正。

【题型 42】反推样本量（误差控制）

- 知识点：CLT+置信要求反解n。
- 经典例题：估计二项比例，要求$P(|\hat{p} - p| \leq 0.03) \geq 0.95$，求最小样本量上界。

解题关键词：最保守取$p(1-p) \leq \frac{1}{4}$，代入$z_{0.975} = 1.96$。

F. 往年高频综合模型（与 6.2 同风格）

【题型 43】圆周三点成钝角三角形

- 知识点：几何概型+等价参数化。
- 经典例题：在圆周随机取三点，求钝角三角形概率。

解题关键词：三段弧长模型，条件$\max\{x,y,z\} > \frac{1}{2}$。

【题型 44】集齐模型（Coupon Collector）

- 知识点：分阶段几何分布求总期望。
- 经典例题：有m种等概率款式，求集齐全部款式所需次数期望。

解题关键词：$E(T) = m\sum_{k=1}^{m}\frac{1}{k}$。

【题型 45】泊松拆分与独立性

- 知识点：Poisson thinning、Y与$X-Y$独立。
- 经典例题：$X \sim P(\lambda)$，每个个体独立保留概率p，令Y为保留数，求Y分布并判断Y与$X-Y$关系。

解题关键词：$Y \sim P(\lambda p), X-Y \sim P(\lambda(1-p))$且独立。

【题型 46】检测阳性后的患病后验概率

- 知识点：低患病率场景下贝叶斯反直觉。
- 经典例题：患病率、灵敏度、假阳率已知，求$P(\text{患病}|\text{阳性})$。

解题关键词：分子“真阳性”，分母“所有阳性”。

【题型 47】抢红包模型的条件分布题

- 知识点：分层条件均匀分布、联合密度与边缘化。
- 经典例题：给定X_1分布及$X_2|X_1$分布，求f_{X_2}并比较$E(X_1),E(X_2)$。

解题关键词：先写$f(x_1,x_2) = f_{X_1}f_{X_2|X_1}$，再按区域积分。

使用建议：

- 第一轮：先做 A+B，保证定义与基本公式熟练。
- 第二轮：主攻 C+D，重点在“联合-边缘-条件-函数分布”链条。
- 第三轮：刷 E+F，训练建模与综合推导速度。

七·数理统计学基本知识

7.1 数理统计学的基本概念

定义 **7.1.1**（统计学基本概念）

- (1)总体：研究对象的全体。
- (2)个体：总体中的单个元素。
- (3)样本：从总体中抽取的一部分个体。
- (4)样本容量：样本中个体的数量。
- (5)实现/样本值：一次实验中样本的观测值。
- (6)简单随机样本：若总体X的样本$(X_1, X_2, ..., X_n)$满足独立同分布。（今后如无特殊说明，样本均指简单随机样本）
- (7)简单随机样本的获得：一般，对有限总体，有放回随机抽样所得到的样本即为简单随机样本；但通常可用不放回随机抽样代替（若满足$\frac{N}{n} \geq 10$）
- (8)统计量：设X总体的样本为$(X_1, X_2, ..., X_n)$，如果$T = g(X_1, X_2, ..., X_n)$是样本的实值连续函数，并且不依赖于未知参数，则称$T$为统计量。若$(x_1, x_2, ..., x_n)$是样本的实现（样本值），则$T(x_1, x_2, ..., x_n)$是统计量的一个观察值。
- (9)统计推断：分为统计估计和假设检验两大类

7.2 总体分布的近似

对于离散型，有频率分布表。设$(x_1, x_2, ..., x_n)$为样本值，设其取到的值为$(a_1, a_2, ..., a_k)$且取到这些值的个数（即频数）为$v_1, v_2, ..., v_k$，易见，$\sum_{i=1}^{k} v_i = n$，从而相应频率为$f_i = \frac{v_i}{n}$，可用如下频率分布表近似给出总体的分布律：

观察值a_k	$a_1 a_2 ... a_k$	合计
频数v_k	$v_1 v_2 ... v_k$	n
频率f_k	$f_1 f_2 ... f_k$	1

对于连续型，有频率分布直方图。设$(x_1, x_2, ..., x_n)$为样本值，设其取到的值的范围为$[a, b]$，将该区间等分成m个小区间$[a, a+\delta), [a+\delta, a+2\delta), ..., [a+(m-1)\delta, b]$，其中$\delta = \frac{b-a}{m}$，设落在第$i$个小区间内的样本值的个数（即频数）为$v_i$，则相应频率为$f_i = \frac{v_i}{n}$，可用频率分布表/频率分布直方图近似给出总体的分布律。

此外，还有经验分布函数。

定义 **7.2.1**（经验分布函数）设X总体的样本为$(X_1, X_2, ..., X_n)$，又设Y_n等可能地取到$(X_1, X_2, ..., X_n)$中的任一个，则称$F_n(x) = P(Y_n \leq x)$为X的经验分布函数。

若样本值的具体取值记作$(x_1, x_2, ..., x_r)$，且其中x_i出现的次数为n_i，则$F_{n(x)} = \sum_{x_i \leq x} \frac{n_i}{n}$。

定理 **7.2.1**（**Glivenko-Cantelli** 定理）设X总体的分布函数为$F(x)$，则对于任意$\varepsilon > 0$，有

$$P\left(\sup_x |F_n(x) - F(x)| > \varepsilon\right) \to 0$$

其中$F_n(x)$为X的经验分布函数。

7.3 常用统计量

定义 7.3.1（几个常用统计量）设X总体的样本为$(X_1, X_2, ..., X_n)$，则

- (1)样本均值：$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- (2)样本方差：$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$
- (3)样本标准差：$S = \sqrt{S^2}$
- (4)样本k阶原点矩：$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$
- (5)样本k阶中心矩：$B_k = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^k$
- (6)顺序统计量与极差：设样本值为$(x_1, x_2, ..., x_n)$，其中$x_1 \le x_2 \le ... \le x_n$，则称$x_i$为第$i$个顺序统计量，$D_n = x_n - x_1$为极差。

至于为何我们抽样可以估计总体，以及为何样本方差的定义是$\frac{1}{n-1}$而非$\frac{1}{n}$，我们给出下面几个结论来说明：

定理 7.3.1（样本的数字特征的几个性质）设$E(X) = \mu, D(X) = \sigma^2$，则

- (1)$E\left(\overline{X}\right) = \mu, D\left(\overline{X}\right) = \frac{\sigma^2}{n}$
- (2)$E\left(S_n^2/B_2\right) = \frac{n-1}{n}\sigma^2$
- (3)$E\left(S^2\right) = \sigma^2$

证：

- (1)

$$E\left(\overline{X}\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \times n \times \mu = \mu$$

$$D\left(\overline{X}\right) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{1}{n^2} \times n \times \sigma^2 = \frac{\sigma^2}{n}$$

前者说明样本均值作为估计量是无偏的，不会系统性地高估或低估；而后者则量化了样本均值的波动大小。n越小，波动越大，单次抽样结果可能离μ很远；n很大时，波动很小，才总会很接近μ。

- (2)

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$$

$$E\left(X_i^2\right) = D(X_i) + E(X_i)^2 = \sigma^2 + \mu^2$$

$$E\left(\overline{X}^2\right) = D\left(\overline{X}\right) + E\left(\overline{X}\right)^2 = \frac{\sigma^2}{n} + \mu^2$$

则

$$E(S_n^2/B_2)=\frac{n-1}{n}\sigma^2$$

这说明直接用n作分母的样本方差系统性地偏小，它是有偏估计。

- (3)由(2)可得

$$E(S^2)=E\left(\frac{n}{n-1}S_n^2\right)=\sigma^2$$

这说明用$n-1$作分母的样本方差是无偏估计，也证明了为什么样本方差定义中要用$n-1$而非n。

7.4 来自正态总体的常用抽样分布

首先介绍$\chi^2(n)$分布（n为自由度）。

定义 7.4.1（χ^2分布）设$X_1,X_2,...,X_n$为$X\sim N(0,1)$的样本，则随机变量$Y=\sum_{i=1}^n X_i^2$的分布称为自由度为n的卡方分布，记为$Y\sim\chi^2(n)$，其概率密度为

$$f(x)=\begin{cases}\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}e^{-\frac{x}{2}} & x>0\\ 0 & x\le 0\end{cases}$$

定理 7.4.1（χ^2分布的几个性质）设$Y\sim\chi^2(n)$，则

- (1)$E(Y)=n,D(Y)=2n$
- (2)若$X_1\sim\chi^2(n_1),X_2\sim\chi^2(n_2),...$，且相互独立，则$\sum_{i=1}^k X_i\sim\chi^2\left(\sum_{i=1}^k n_i\right)$
- (3)当n较大时，Y近似服从$N(n,2n)$。

证：

- (1)由定义，$Y=\sum_{i=1}^n X_i^2$，又$E(X_i^2)=D(X_i)+E(X_i)^2=1$，故$E(Y)=n$；又$D(X_i^2)=E(X_i^4)-E(X_i^2)^2=3-1=2$，且$X_i^2$相互独立，故$D(Y)=\sum_{i=1}^n D(X_i^2)=2n$。

下面介绍$t(n)$分布（n为自由度）。

定义 7.4.2（t分布）设$X\sim N(0,1),Y\sim\chi^2(n)$且相互独立，则随机变量$T=\frac{X}{\sqrt{\frac{Y}{n}}}$的分布称为自由度为$n$的$t$分布，记为$T\sim t(n)$，其概率密度为

$$f(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}},x\in R$$

定理 7.4.2（t分布的几个性质）设$T\sim t(n)$，则

- (1)$f_n(t)$是偶函数，且在$t=0$处取得最大值。
- (2)$n\to+\infty$时，$f_n(t)\to\Phi(t)$，其中$\Phi(t)$是标准正态分布的密度函数。

最后介绍$F(n,m)$分布（n,m为第一、二自由度）。

定义 7.4.3（F分布）设$X \sim \chi^2(n), Y \sim \chi^2(m)$且相互独立，则随机变量$F = \frac{\frac{X}{n}}{\frac{Y}{m}}$的分布称为第一、二自由度分别为$n, m$的$F$分布，记为$F \sim F(n, m)$，其概率密度为

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)}\left(\frac{n}{m}\right)^{\frac{n}{2}} x^{\frac{n}{2}-1}\left(1+\frac{n}{m}x\right)^{-\frac{n+m}{2}} & x > 0 \\ 0 \quad x \le 0 \end{cases}$$

定理 7.4.3（F分布的几个性质）设$F \sim F(n, m)$，则

- (1)若$X \sim F(n, m)$，则$\frac{1}{X} \sim F(m, n)$
- (2)$F_{1-\alpha}(n, m) = \frac{1}{F_\alpha(m, n)}$

证：

- (1)由定义，$F = \frac{\frac{X}{n}}{\frac{Y}{m}}$，则$\frac{1}{F} = \frac{\frac{Y}{m}}{\frac{X}{n}}$，又$Y \sim \chi^2(m), X \sim \chi^2(n)$且相互独立，故$\frac{1}{F} \sim F(m, n)$。
- (2)由(1)，$P(X > F_{1-\alpha}(n, m)) = P\left(\frac{1}{X} < \frac{1}{F_{1-\alpha}(n,m)}\right) = P\left(\frac{1}{X} < F_\alpha(m, n)\right) = \alpha$，又$P(X > F_\alpha(n, m)) = \alpha$，故$F_{1-\alpha}(n, m) = \frac{1}{F_\alpha(m,n)}$。

【例】证明：$t^2_{1-\frac{\alpha}{2}}(n) = F_\alpha(1, n)$

证：由题得$\begin{cases} 0 \le 1-\frac{\alpha}{2} \le 1 \\ 0 \le \alpha \le 1 \end{cases}$即$\alpha \in [0, 1]$

由定义，$t^2(n) = \frac{\frac{X^2}{1}}{\frac{Y}{n}}$，又$X^2 \sim \chi^2(1), Y \sim \chi^2(n)$且相互独立，设$W \sim t(n)$，令$Z = W^2$，则$Z \sim F(1, n)$，且$P\left(W < t_{1-\frac{\alpha}{2}}(n)\right) = \frac{\alpha}{2}$，故$P(Z < F_\alpha(1, n)) = 1 - \frac{\alpha}{2}$，即$t^2_{1-\frac{\alpha}{2}}(n) = F_\alpha(1, n)$。

下面我们介绍正态总体之样本均值与样本方差的一些结论：

定理 7.4.4（正态总体的样本均值与样本方差的分布）设$X \sim N(\mu, \sigma^2)$，样本为$(X_1, X_2, ..., X_n)$，则

- (1)样本均值$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
- (2)样本方差$S^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1)$
- (3)样本均值与样本方差相互独立
- (4)$\frac{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\frac{S}{\sigma}} \sim t(n-1)$

若$(X_1, X_2, ..., X_n)$与$(Y_1, Y_2, ..., Y_m)$分别是$X \sim N(\mu_1, \sigma_1^2)$与$Y \sim N(\mu_2, \sigma_2^2)$的样本，且相互独立，则

- (5)$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F(n-1, m-1)$

若$\sigma_1 = \sigma_2$，则

- (6)$\frac{S_1^2}{S_2^2} \sim F(n-1, m-1)$
- (7)$\frac{\left(\overline{X}-\overline{Y}\right)-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}} \sim t(n+m-2)$

证：

- (1)由定义，$\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$，又$X_i \sim N(\mu, \sigma^2)$且相互独立，故$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$。

- (2)由定义, $S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$, 又$X_i\sim N(\mu,\sigma^2)$且相互独立, 故$S^2\sim\frac{\sigma^2}{n-1}\chi^2(n-1)$。
- (3)由定义, $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$, 又$S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$, 则$\overline{X}$与$S^2$是两个不同的二次型, 且它们的矩阵满足$AB=0$, 故它们相互独立。
- (4)由(1)(2)(3), $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$, $S^2\sim\frac{\sigma^2}{n-1}\chi^2(n-1)$且相互独立, 故$\frac{\frac{\frac{\overline{X}-\mu}{\sigma}}{\sqrt{n}}}{\frac{S}{\sigma}}\sim t(n-1)$。
- (5)由定义, $\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}=\frac{\left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\frac{\left(X_i-\overline{X}\right)^2}{\sigma_1^2}}{\left(\frac{1}{m-1}\right)\sum_{j=1}^{m}\frac{\left(Y_j-\overline{Y}\right)^2}{\sigma_2^2}}$, 又$S_1^2\sim\frac{\sigma_1^2}{n-1}\chi^2(n-1), S_2^2\sim\frac{\sigma_2^2}{m-1}\chi^2(m-1)$且相互独立, 故$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}\sim F(n-1,m-1)$。
- (6)由(5), $\frac{S_1^2}{S_2^2}=\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}\sim F(n-1,m-1)$。
- (7)由 (1)(2)(3)(6), $\left(\overline{X}-\overline{Y}\right)-(\mu_1-\mu_2)\sim N\left(0,\left(\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}\right)\right)$, $\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}\sim\frac{(n-1)\sigma_1^2+(m-1)\sigma_2^2}{n+m-2}\chi^2(n+m-2)$且相互独立, 故$\frac{\left(\overline{X}-\overline{Y}\right)-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}}\sim t(n+m-2)$。