第 4 讲　常见随机变量的概率分布

知识梳理

一　常见随机变量分布一览

1. 常见一元离散分布

分布名称	符号	分布律	期望	方差	说明
0–1 分布	$X \sim 0-1(p)$	$P(X=0)=1-p \quad P(X=1)=p$	p	$p(1-p)$	
二项分布	$X \sim B(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k}$	np	$np(1-p)$	不直说
泊松分布	$X \sim \pi(\lambda)$	$P(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda}$	λ	λ	记前几个

2. 常见一元连续分布

分布名称	符号	密度函数	期望	方差	
均匀分布	$X \sim U(a,b)$	$f(x)=\begin{cases}1/(b-a), & x\in(a,b)\\ 0, & \text{其它}\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	——
正态分布	$X \sim N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty$	μ	σ^2	化成标准
指数分布	$X \sim E(\lambda)$	$f(x)=\begin{cases}\lambda e^{-\lambda x}, x>0\\ 0, \quad x\le 0\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	无记忆性 记住分布函数

3. 常见二元连续分布

分布名称	符号 或 密度函数	备注
二元均匀分布	$f(x,y)=\begin{cases}c, & (x,y)\in D\\ 0, & \text{其它}\end{cases}$	c 为区域 D 面积的倒数
二元正态分布	$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$	ρ 就是 X 和 Y 的相关系数

二　随机变量数字特征的性质

1. 数学期望的性质

数学期望的线性组合（Xi 无需相互独立）

$$E(c_0+\sum_{i=1}^n c_iX_i)=c_0+\sum_{i=1}^n c_iE(X_i)$$

数学期望的乘积（Xi 必须相互独立）

$$E(\prod_{i=1}^n X_i)=\prod_{i=1}^n E(X_i)$$

2. 方差的性质

方差的线性组合（Xi 无需相互独立）

$$\mathrm{Var}(\sum_{i=1}^n c_iX_i+c_0)=\sum_{i=1}^n c_i^2\mathrm{Var}(X_i)+2\sum_{1\le i<j\le n}\mathrm{Cov}(X_i,X_j)$$

· 当 X_i 两两独立时，协方差一项为 0

3. **协方差的性质**

协方差的性质 1	协方差的性质 2	协方差的性质 3
$\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$	$\mathrm{Cov}(X,X)=\mathrm{Var}(X)$	$\mathrm{Cov}(aX,bY)=ab\mathrm{Cov}(X,Y)$

协方差的性质 1	协方差的性质 2
$\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$	若 X 和 Y 相互独立，则 $\mathrm{Cov}(X,Y)=0$

三　正态分布的重要性质

1. **一元正态分布的性质**

① **标准正态分布 $Z\sim N(0,1)$ 的分布函数 $\Phi(x)$**

标准正态分布函数	标准正态分布函数的性质	标准正态分布函数的性质
$\Phi(x)=P(Z\le x)$	$\Phi(-x)=1-\Phi(x)$	$P(a<Z<b)=\Phi(b)-\Phi(a)$

· $\Phi(x)$ 的值通过手册查阅，**考试会提供需要用到的 $\Phi(x)$ 值**

② **正态分布的线性函数依然是正态分布** → 可经线性函数转化为标准正态分布

正态分布 → 标准正态分布的转化
$X\sim N(\mu,\sigma^2) \;\rightarrow\; Z=\dfrac{X-\mu}{\sigma}\sim N(0,1)$

· 因此求解正态分布问题，往往需要转化为标准正态分布处理

2. **二元正态分布的性质**

① **二元正态分布的 X 和 Y 各自依然遵循正态分布**

二元正态分布的分解
$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho) \;\rightarrow\; X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$

② **X 和 Y 的线性函数 $aX+bY+c$ 依然是正态分布**

· 此时新正态变量的 μ 和 σ^2 分别是 $aX+bY+c$ 的均值和方差

③ **对于两个正态变量，不相关和独立是等价的**

· 这使得判断正态变量是否独立时，只需计算协方差即可

四　其他分布的重要性质

1. **指数分布的性质**

指数分布的无记忆性
$P(X>a+b\mid X>a)=P(X>b)$

题型解析

八　求解常见随机变量分布

1. 题型简述与解法

- 通常以填空题的形式出现（偶尔也会有大题），考查常见随机变量分布的事件概率、期望、方差等
- 解题关键在于记住这些分布的期望、方差以及出众的特征，可以大幅减少时间和脑力消耗
- 与一般的分布不同，常见随机变量分布往往是通过 $D(X)$ 反求 $E(X^2)$，需要注意

① **0-1 分布**　例 1

- 该分布比较简单，只需要能反应过来它只有两种取值就可以了

② **二项分布**

- 该分布通常不会在题干中明确指出，往往通过“对 xx 进行独立考察”等字眼暗示
- 此时独立考察的次数就是参数 n， p 则要根据题意自己求
- 通常会求事件概率，枚举复杂时可尝试求逆事件

③ **泊松分布**

- 该分布通常会求事件概率，枚举复杂时可尝试求逆事件
- 最好记住前 3 个取值的分布律，加快解题速度：

$$P(X=0)=\mathrm{e}^{-\lambda}\,,\quad P(X=1)=\lambda\mathrm{e}^{-\lambda}\,,\quad P(X=2)=\frac{\lambda^2}{2}\mathrm{e}^{-\lambda}$$

④ **均匀分布**

- 最好记住均值和方差，加快解题速度

⑤ **正态分布**

- 求事件概率要将正态分布 X 转换为标准正态分布 Z，注意要对事件作一样的转换

$$P(x_1<X<x_2)=P(\frac{x_1-\mu}{\sigma}<Z<\frac{x_2-\mu}{\sigma})$$

- 如果试卷开头找不到需要的 Φ，那你肯定算错了

⑥ **指数分布**

- 记住分布函数以及事件概率

$$F(x)=P(X\le x)=1-\mathrm{e}^{-\lambda x}\qquad P(X>x)=\mathrm{e}^{-\lambda x}$$

- 记住指数分布的无记忆性，可以将条件概率直接转为更简单的概率

⑦ **二元均匀分布**

- 求事件概率计算面积乘上概率密度函数即可，尤其是区域为圆的时候

⑧ **二元正态分布**

- 一般考察正态变量之间的相关性，新正态变量的期望与方差等

2.历年考试典型例题

例 1 （**19－20 春夏**）设 X 与 Y 均服从 0－1 分布，$P(X=1)=\frac{3}{4}$，$P(X=1,Y=1)=\frac{1}{4}$

（1）若 X 与 Y 独立，则 $P(X=0,Y=1)=$________；

（2）若 X 与 Y 的协方差为 $-\frac{1}{16}$，则 $P(X=0,Y=1)=$________.

解 因为 X 与 Y 均服从 0－1 分布，可直接打表，并得到 $P(X=1,Y=0)=1/2$

$X\backslash Y$	0	1	
0		?	1/4
1	1/2	1/4	3/4
	1－p	p	

（1）若 X 与 Y 独立，则 $P(X=1,Y=0)=P(X=1)P(Y=0)=1/2 \rightarrow P(Y=0)=2/3$

因此 $P(Y=1)=1/3$，$P(X=0,Y=1)=P(X=0)P(Y=1)=\boxed{1/12}$

（2）若 X 与 Y 的协方差为 $-\frac{1}{16}$，由 $E(X)=3/4$，$E(Y)=p$，$E(XY)=P(X=1,Y=1)=1/4$

即 $\mathrm{Cov}(X,Y)=\frac{1}{4}-\frac{3}{4}p=-\frac{1}{16}$，解得 $p=\frac{5}{12}$，因此 $P(X=1,Y=0)=\frac{5}{12}-\frac{1}{4}=\boxed{1/6}$

例 2 （**14－15 春夏 节选**）某人喜欢长跑，基本上每天跑 10 公里，假设他跑 10 公里所花时间（分钟）为 $Y=40+20X$，其中 X 的概率密度为 $f(x)=\begin{cases}2(1-x),0<x<1\\ 0, \quad 其他\end{cases}$.

（3）若一周跑 6 次，每天用时是相互独立的，求至少有 5 次用时少于 50 分钟的概率；

解 原题的前两问已经求出 $P(Y<50)$ 的概率，此处不再赘述过程，结果为 0.75

由题意可得这是一个二项分布，$n=6$，$p=0.75$

设少于 50 分钟的次数为 Z，则 $P(Z=k)=C_6^k(\frac{3}{4})^k(\frac{1}{4})^{6-k}$ （$k=0,\cdots,6$）

因此题目要求的 $P(Z\geq 5)=P(Z=5)+P(Z=6)=C_6^5(\frac{3}{4})^5\frac{1}{4}+C_6^6(\frac{3}{4})^6=\boxed{0.534}$

例 3 某公交车站单位时间内等车的人数 X 服从泊松分布 $\pi(\lambda)$，$\lambda>0$

（1）（**18－19 春夏**）若 $E(X^2)=2\mathrm{Var}(X)$，则 $\lambda=$________.

（2）（**17－18 秋冬**）X 的分布函数值 $F(2)=$______.

（3）（**15－16 秋冬**）若已知有人在等车，则至少有 2 人等车的概率为________.

解　（1）由 $X\sim\pi(\lambda)$，得 $E(X)=\mathrm{Var}(X)=\lambda$，则 $E(X^2)=D(X)+[E(X)]^2=\lambda+\lambda^2$

因此 $E(X^2)=2\mathrm{Var}(X)\ \to\ \lambda+\lambda^2=2\lambda$，$\lambda>0\ \to\ \boxed{\lambda=1}$

（2）$F(2)=P(X\le 2)=P(X=0)+P(X=1)+P(X=2)=(1+\lambda+\dfrac{\lambda^2}{2})e^{-\lambda}=\boxed{2.5e^{-1}}$

（3）已知有人在等车 $\to\ X>0$，因此所求概率为

$$P(X\ge 2\,|\,X>0)=\frac{P(X\ge 2,X>0)}{P(X>0)}=\frac{1-P(X=1)-P(X=0)}{1-P(X=0)}=\frac{1-e^{-1}-e^{-1}}{1-e^{-1}}=\boxed{\frac{1-2e^{-1}}{1-e^{-1}}}$$

例 4　（1）（**17—18 春夏**）设 $X\sim U(1,c)$（均匀分布），$E(X)=2$，则 $c=$____，$\mathrm{Var}(X)=$_____.

（2）（**18—19 秋冬**）设 (X,Y) 在区域 $\{(x,y):0\le x\le 1,0\le y\le 1\}$ 上均匀分布，则

$P(X^2+Y^2\le 1)=$___________.

解　（1）由 $E(X)=(1+c)/2=2$，解得 $c=3$，$\mathrm{Var}(X)=(3-1)^2/12=1/3$

（2）总的区域为 1×1 的正方形，要求的区域为四分之一圆，因此概率为 $(\pi/4)/1=\pi/4$

例 5　（**16—17 秋冬**）设 $(X,Y)\sim N(1,0,1,4,0.76)$，则 $P(X<Y-0.4)=$_________；当 $c=$_______ 时，$cX-Y$ 与 $X-Y$ 相互独立.

解　由 $(X,Y)\sim N(1,0,1,4,0.76)$，得 $X\sim N(1,1)$，$Y\sim N(0,4)$，X 与 Y 的相关系数为 0.76（注意参数顺序是 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$），则

$$E(X-Y)=1-0=1$$

$$D(X-Y)=D(X)+D(Y)-2\rho\sqrt{D(X)D(Y)}=1.96=1.4^2$$

因此 $X-Y\sim N(1,1.4^2)$：

$$P(X<Y-0.4)=P(X-Y<-0.4)=P(\frac{X-Y-1}{1.4}<\frac{-0.4-1}{1.4})=\Phi(-1)=1-\Phi(1)=\boxed{0.16}$$

由于 $\mathrm{Cov}(cX-Y,X-Y)=cD(X)-(c+1)\mathrm{Cov}(X,Y)+D(Y)=c-1.52(c+1)+4$

由正态分布的性质，当 $c-1.52(c+1)+4=0$ 时，相互独立 $\to\ \boxed{c=4.77}$

例 6　（**14—15 春夏**）某小店开门到首个顾客到达所用的时间 X（单位：分钟）的概率密度函数为

$$f(x)=\begin{cases}0.2e^{-0.2x}, & x>0\\ 0, & x\le 0\end{cases}$$

则在开门后 10 分钟内首个顾客到达的概率为___________，若开门后 5 分钟内没有顾客到达，则在接下来的 5 分钟内仍没有顾客到达的概率为____________.

解　由题意 $X\sim E(0.2)$，即 X 服从 $\lambda=0.2$ 的指数分布，因此第一空 $P(X<10)=1-e^{-0.2\times 10}=1-e^{-2}$

第二空 $P(X>5+5\,|\,X>5)=P(X>5)=1-e^{-0.2\times 5}=1-e^{-1}$