考研数学概率论与数理统计强化讲义

主讲：张宇

张宇：新东方在线名师，博士，全国著名考研数学辅导专家，教育部“国家精品课程建设骨干教师”，全国畅销书《高等数学 18 讲》、《考研数学题源探析经典 1000 题》作者，高等教育出版社《全国硕士研究生入学统一考试数学考试参考书（大纲解析）》编者之一，2007 年斯洛文尼亚全球可持续发展大会受邀专家（发表 15 分钟主旨演讲）. 首创“题源教学法”，对考研数学的知识结构和体系有全新的解读，对考研数学的命题与复习思路有极强的把握和预测能力，让学生轻松高效夺取高分.

欢迎使用新东方在线电子教材

目　录

第一讲 随机事件与概率

综述

1．正确理解概念，灵活使用公式

2．用古典概型、几何概型、公式求复杂事件的概率

一、重要概念与公式

1．事件的关系与运算

（1）样本空间——Ω（全集）

其基本元素 w_i，叫样本点.

（2）事件——样本空间的子集：A、B、C、$\cdots$

$\varnothing$——不可能事件

Ω——必然事件

（3）完备事件组

$\bigcup\limits_i A_i=\Omega$，$A_iA_j=\varnothing$（$i\neq j$）

（4）运算关系

$\overline{A\cup B}=\overline{A}\cap\overline{B}$，$\overline{A\cap B}=\overline{A}\cup\overline{B}$

2．古典概型

若 Ω 中有有限个、等可能的样本点，称为古典概型.

$$P(A)=\frac{A\text{中样本点个数}}{\Omega\text{中样本点个数}}$$

【例】设 5 封信投入 4 个信箱.

求下列事件的概率：

A_1={1、2 号信箱中各有 1 封信}

A_2={某信箱中有 3 封信}

A_3={第 2 个信箱没有信}

A_4={仅有一个信箱没有信}

3．几何概型

（1）引例

（2）定义

若Ω是一个可度量的几何区域，且样本点落入Ω中的某一可度量子区域A的可能性大小与A的几何度量成正比，而与A的位置与形状无关，称为几何概型.

$$P(A)=\frac{\text{A的度量}}{\Omega\text{的度量}}$$

【例】在$(0,1)$内随机取两个数，则这两个数之差的绝对值小于$\frac{1}{2}$的概率为________.

4．重要公式

（1）对立 $P(A)=1-P(\overline{A})$

（2）减法 $P(A-B)=P(A)-P(AB)$

（3）加法

1）$P(A+B)=P(A)+P(B)-P(AB)$

2） $P(A+B+C)=P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC)$

【注】超过三个的事件和的概率，一般附加“互斥”、“独立”条件.

①若 A_1， A_2， $\cdots$， A_n $(n>3)$ 两两互斥，则 $P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)$.

②设 A_1， A_2， $\cdots$， A_n.

若对其中任意有限个 A_{i1}， A_{i2}， $\cdots$， $A_{ik}(k\geq 2)$，都有

$P(A_{i1}A_{i2}\cdots A_{ik})=P(A_{i1})P(A_{i2})\cdots P(A_{ik})$，则称 A_1， A_2， $\cdots$， A_n 相互独立.

$$\begin{cases} P(A_1A_2)=P(A_1)P(A_2) \\ P(A_1A_3)=P(A_1)P(A_3) \\ P(A_2A_3)=P(A_2)P(A_3) \\ P(A_1A_2A_3)=P(A_1)P(A_2)P(A_3) \end{cases} \Leftrightarrow A_1,A_2,A_3 \text{相互独立}.$$

n 个事件相互独立 $\Leftrightarrow$ 它们中任意一部分事件换成各自的对立事件所得 n 个新事件相互独立.

若 A_1， A_2， $\cdots$， A_n $(n>3)$ 相互独立，则

$$P(\bigcup_{i=1}^{n}A_i)=1-P(\overline{\bigcup_{i=1}^{n}A_i})=1-P(\bigcap_{i=1}^{n}\overline{A_i})=1-\prod_{i=1}^{n}P(\overline{A_i})=1-\prod_{i=1}^{n}[1-P(A_i)]$$

（4）条件 $P(A|B)=\dfrac{P(AB)}{P(B)}$， $P(B)>0$

（5）乘法 $P(AB)=P(A)P(B|A)=P(B)P(A|B)$

$$P(A_1A_2A_3)=P(A_1A_2)P(A_3|A_1A_2)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)$$

（6）全集分解公式（全概公式）

1）引例

一个村子 三个小偷

2）若试验分成两阶段

（Ⅰ）选人

（Ⅱ）去偷

如果$\bigcup_{i-1}^{n}A_i=\Omega, A_iA_j=\phi(i\neq j), P(A_i)>0$，则对任一事件 B，有 $B=\bigcup_{i-1}^{n}A_iB$，

$$P(B)=\sum_{i-1}^{n}P(A_i)P(B\mid A_i).$$

（7）贝叶斯公式（逆概公式）

如果 $\bigcup_{i=1}^{n}A_i=\Omega, A_iA_j=\phi(i\neq j), P(A_i)>0$，则对任一件事 B，只要 $P(B)>0$，有

$$P(A_i\mid B)=\frac{P(A_i)P(B\mid A_i)}{\sum_{i=1}^{n}P(A_i)(B\mid A_i)}(i=1,2,\cdots,n)$$

【例 1】下列说法，正确的是（ ）

（A）已知 $P(A)=P(B)=\frac{1}{2}$， $P(A+B)=1$，则 $\overline{A}\cup\overline{B}=\Omega$

（B）设 $P(A)>0$， $P(B)>0$，则“A、B 互斥”与“A、B 独立”可同时成立

（C）将一枚硬币独立掷两次，记 A_1={第一次正面}， A_2={第二次正面}， A_3={正、反面各一次}，则 A_1， A_2， A_3 相互独立.

（D）袋中有 100 个球，40 白 60 黑，从袋中先后不放回取 100 次，则第 100 次取到白的概率为 $\frac{2}{5}$

【例 2】以下结论，错误的是（ ）

（A）任意事件 A、B、C 均有 $P(AB)+P(AC)+P(BC)\geq P(A)+P(B)+P(C)-1$

（B）若 $P(A|B)>P(A|\overline{B})$，则 $P(B|A)>P(B|\overline{A})$

（C）任意事件 A_1，A_2，$\cdots$，A_n，均有

$P(A_1A_2\cdots A_n)\geq P(A_1)+P(A_2)+\cdots+P(A_n)-(n-1)$

（D）任意事件 A、B 均有 $|P(AB)-P(A)P(B)|>\dfrac{1}{4}$

【例 3】在大街上随机采访 6 个人，求下列事件的概率：

（1）至少一人在 6 月份出生；

（2）恰有 4 人在第四个季度出生；

（3）恰有 3 个人在同一月出生.

【例 4】将 4 位考生的录取通知书随机装入 4 个印有他们名字的信封，求 4 封通知书全装错的概率.

【例 5】设有 10 份报名表，3 女 7 男.现从中每次取一份，取后不放回，求下列事件的概率：
（1）第 3 次取到女；（2）第 3 次才取到女；（3）已知前两次没取到女，第三次取到女.

【例 6】设有三个地区各 10、15、25 名考生的报名表，其中女生表分别为 3、7、5 份.随机取一个地区的报名表，从中先后无放回取两份.
（Ⅰ）求先取到的一份为女的概率；
（Ⅱ）已知后取到的一份是男，求先取到的是女的概率.

【例 7】有两批数量相同的灯泡，已知有一批全正品，另一批有 $\frac{1}{4}$ 次品，$\frac{3}{4}$ 正品.现从两批产品中任取一个灯泡，经检验为正品，放回原处并在原所在批次再取一个灯泡，求此灯泡是次品的概率.

第二讲 一维随机变量及其分布

综述

1．八个重要分布

2．一维随机变量 X 与其分布函数 $F(x)$

3．随机变量函数 $Y=g(X)$ 的分布 $F(y)$

一、概念与八个分布

1．随机变量 X 与分布函数 $F(x)$

①随机变量

随机变量就是“其值随机会而定”的变量．设随机试验 E 的样本空间为 $\Omega=\{\omega\}$，如果对每一个 $\omega\in\Omega$，都有唯一的实数 $X(\omega)$ 与之对应，并且对任意实数 x，$\{\omega: X(\omega)\le x\}$ 是随机事件，则称定义在 Ω 上的实单值函数 $X(\omega)$ 为随机变量．简记为随机变量 X．一般用大写字母 $X,Y,Z,\cdots$ 或希腊字母 $\xi,\eta,\cdots$ 来表示随机变量.

②分布函数

分布函数 $F(x)\triangleq P\{X\le x\},-\infty<x<+\infty$

2．离散型随机变量

①定义

如果随机变量 X 只可能取有限个或可列个值 $x_1,x_2,\cdots$，则称 X 为离散型随机变量.

②分布律 $X\sim\begin{pmatrix} x_1 & x_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$

③ $F(x)\triangleq P\{X\le x\},-\infty<x<+\infty$

步步高的阶梯型函数

3．连续型随机变量

若存在非负可积函数 $f(x)$，使得任给 $x\in(-\infty,+\infty)$ 有 $F(x)=\int_{-\infty}^{x} f(t)dt$，则称 X 为连续型随机变量，$f(x)$ 称为 X 的概率密度函数.

4． $X \sim F(x)\begin{cases} p_i \\ f(x) \end{cases}$

①$F(x)$ 是某个随机变量 X 的分布函数 $\Leftrightarrow \begin{cases} 1.\text{单调不减} \\ 2.F(-\infty)=0, F(+\infty)=1 \\ 3.\text{右连续} \end{cases}$

②$\{p_i\}$ 是分布律 $\Leftrightarrow \begin{cases} 1.p_i \geq 0 \\ 2.\sum\limits_i p_i = 1 \end{cases}$

③ $f(x)$ 是概率密度函数 $\Leftrightarrow \begin{cases} 1.f(x) \geq 0 \\ 2.\int_{-\infty}^{+\infty} f(x)dx = 1 \end{cases}$

5．八个分布

①$0-1$分布 $B(1,p)$

如果 X 的概率分布为 $X\sim\begin{pmatrix} 1 & 0 \\ P & 1-P \end{pmatrix}$ 即 $P(X=1)=p$， $P(X=0)=1-p$，则称 X 服从参数为 p 的$0-1$分布，记为 $X\sim B(1,p)(0<p<1)$.

②二项分布 $X\sim B(n,p)\begin{cases} 1.\text{独立} \\ 2.P(A)=p \\ 3.\text{只有}A,\overline{A} \end{cases}$

如果 X 的概率分布为 $p_k=P(X=k)=C_n^k p^k(1-p)^{n-k}, k=0,1,\cdots,n, 0<p<1$，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 $X\sim G(p)$　　首中即停止

如果 X 的概率分布为 $p_k = P(X=k) = q^{k-1}p, k=1,2,\cdots$， $0<p<1$， $q=1-p$，则称 X 服从参数为 p 的几何分布，记为 $X \sim G(p)$.

④超几何分布

N 件产品，M 件正品，无放回取 n 次，则取到 k 个正品的概率

$$P(X=k) = \frac{C_M^k C_{N-M}^{n-k}}{C_M^n}, k=0,1,2,\cdots,n$$

⑤泊松分布 $P(\lambda)$

某时间段，某场合下，源源不断的质点来流的个数

$$P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}, k=0,1,2,\cdots$$

λ——强度，$EX=\lambda$

⑥均匀分布——“几何概型”

若 $X \sim f(x) = \begin{cases} \dfrac{1}{b-a}, & a<x<b, \\ 0, & \text{其他} \end{cases}$，则 $X \sim U(a,b)$.

【注】高档次说法：“X 在 I 上的任一子区间取值的概率与该子区间长度成正比”

$\Rightarrow X \sim U(I)$

⑦指数分布

若 $X \sim f(x) = \begin{cases} \lambda e^{-\lambda x}, & x>0, \\ 0, & x \le 0 \end{cases}$，则 $X \sim E(\lambda)(\lambda>0)$.

λ——失效率，$E(X) = \dfrac{1}{\lambda}$

【注】1）$P(X \ge t+s \mid X \ge t) = P(X \ge s)$——无记忆性

2）$F(x) = \begin{cases} 1-e^{-\lambda x}, & x>0, \\ 0, & x \le 0 \end{cases}$

3）$\begin{cases} \text{几何分布} \Leftrightarrow \text{离散型等待分布} \\ \text{指数分布} \Leftrightarrow \text{连续型等待分布} \end{cases} \Leftrightarrow$ 无记忆性

⑧正态分布

若 $X \sim f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，则 $X \sim N(\mu,\sigma^2)$.

$\Phi(-x)=1-\Phi(x)$，$\Phi(0)=\dfrac{1}{2}$.

特别地，当 $\mu=0$，$\sigma^2=1$ 时，$X \sim N(0,1)$.

$$\begin{cases} X \sim \varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},-\infty<x<+\infty \\ X \sim \Phi(x)=\displaystyle\int_{-\infty}^{x}\dfrac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt,-\infty<x<+\infty \end{cases}$$

二、综合题分析

（1）概念

（2）求随机变量的分布函数

（3）求随机变量函数的分布函数

【例 1】下列说法， 错误的是（ ）

（A）随机变量 X_1、X_2 相互独立，$X_1 \sim F_1(x)$，$X_2 \sim F_2(x)$，则 $F_1(x)F_2(x)$ 必为 $\max\{X_1,X_2\}$ 的分布函数

（B）若 $X \sim f(x)=Ae^{-(\frac{x+1}{2})^2}$，则 $A=\dfrac{1}{2\sqrt{\pi}}$

（C）若 $X \sim f(x)=\begin{cases}\dfrac{1}{3}, & 0\le x\le 1, \\ \dfrac{2}{9}, & 3\le x\le 6, \\ 0, & \text{其他}\end{cases}$，且 $P\{X\le k\}=\dfrac{2}{3}$，则 $k<1$

（D）若 $X \sim F(x)$，$X \sim f(x)$，且 $x\le 0$ 时，$f(x)$ 连续，$f(x)=F(x)$，$F(0)=1$，

则 $f(x)=\begin{cases}e^x, & x<0, \\ 0, & x\ge 0\end{cases}$

【例 2】设$\left|X\right|\le 1$，$P(X=-1)=\dfrac{1}{8}$，$P(X=1)=\dfrac{1}{4}$.在$\left\{-1<X<1\right\}$发生的条件下，X在$(-1,1)$内任一子区间取值的条件概率与该子区间长度成正比，求X的分布函数$F(x)$.

【例 3】设一机器在任何长为t的时间内出故障的次数$N(t)$服从参数为λt的泊松分布.

（Ⅰ）求相继两次故障之间的时间间隔T的分布函数$F_T(t)$；

（Ⅱ）求在设备已无故障工作 8 小时的情形下，再无故障工作 8 小时的概率.

【例 4】若 $X \sim f_X(x)$，$Y = g(X)$，求 $Y \sim f_Y(y)$.

【例 5】设 $X \sim f_X(x)=\begin{cases} \frac{1}{2}, & -1<x<0, \\ \frac{1}{4}, & 0 \le x<2, \\ 0, & \text{其他} \end{cases}$，令 $Y = X^2$，求 $Y \sim f_Y(y)$.

第三讲　多维随机变量及其分布

综述

1．概念

2．用分布求概率

3． $Z=g(X,Y)$，求 Z 的分布

一、概念

1．联合分布

设 (X,Y) 为二维随机变量，对任意的实数 x， y，称二元函数

$$F(x,y)\triangleq P\left\{X\le x,Y\le y\right\}，\ -\infty<x<+\infty，\ -\infty<y<+\infty$$

为二维随机变量 (X,Y) 的联合分布函数，简称为分布函数，记为 $(X,Y)\sim F(x,y)$.

2．边缘分布

已知 $F(x,y)$，

$$F_X(x)\triangleq P\left\{X\le x\right\}=P\left\{X\le x,\Omega\right\}=P\left\{X\le x,Y\le+\infty\right\}=F(x,+\infty)=\lim_{y\to+\infty}F(x,y)$$

$$F_Y(y)=\lim_{x\to+\infty}F(x,y)$$

【注】①离散型 $(X,Y)\sim p_{ij}$（联合分布律）

$X \backslash Y$	y_1	$\cdots$	y_j	$\cdots$	$P(X=x_i)$
x_1	p_{11}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
x_i	p_{i1}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
$P(Y=y_j)$	$p_{\cdot 1}$	$\cdots$	$p_{\cdot j}$	$\cdots$	1

$$P(X=x_i,Y=y_j)=p_{ij}$$

X 、Y 的边缘分布分别为

$$p_i\cdot \hat{=} P\{X=x_i\}=\sum_j P\{X=x_i,Y=y_j\}=\sum_j p_{ij}(i=1,2,\cdots)$$

$$p\cdot_j \hat{=} P\{Y=y_i\}=\sum_i p_{ij}(j=1,2,\cdots)$$

其“条件分布”为

$$P\left\{X=x_i \middle| Y=y_j\right\}=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{p_{\cdot j}}$$ 条件=联合/边缘

②连续型$(X,Y)\sim f(x,y)$ 联合概率密度

边缘概率密度

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

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

其条件概率密度为

$$f_{X|Y}(x\,|\,y)=\frac{f(x,y)}{f_Y(y)}(f_Y(y)>0).$$

3．独立性

X,Y 独立$\Leftrightarrow F(x,y)=F_X(x)\cdot F_Y(y)$

$\Leftrightarrow p_{ij}=p_{i\cdot}\cdot p_{\cdot j}$，任给$i$， j

$$\Leftrightarrow f(x,y)=f_X(x)\cdot f_Y(y)$$

4．两大分布

①均匀分布

$$(X,Y)\sim f(x,y)=\begin{cases}\dfrac{1}{S_D}, & (x,y)\in D\\ 0, & (x,y)\notin D\end{cases}$$

②正态分布

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

1^0 X,Y 的边缘分布一定是正态分布；

2^0 $aX+bY$ 仍然服从正态分布；

3^0 独立 $\Leftrightarrow$ 不相关.

2）若 $X_1,X_2,\cdots,X_n$ 相互独立，且均服从正态分布，则 $X_1,X_2,\cdots,X_n$ 的线性组合仍然服从正态分布.

二、综合题解析

【例 1】一设备由两部件组成，以 X 、Y 分别表示两部件寿命（单位：千小时）.

$$(X,Y)\sim F(x,y)=\begin{cases}1-e^{-0.5x}-e^{-0.5y}+e^{-0.5(x+y)}, & x\geq 0,y\geq 0,\\ 0, & \text{其他}\end{cases}$$

问 X 、Y 是否独立？

【例 2】设 $X_1 \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$，$X_2 \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$，且 $P(X_1X_2=0)=1$.

（Ⅰ）求 X_1、X_2 的联合分布律 p_{ij}；（Ⅱ）X_1、X_2 独立吗？

【例 3】设 $X \sim U(0,1)$，在 $X \sim x(0<x<1)$ 的条件下，Y 在 $(0,x)$ 内服从均匀分布，求

（Ⅰ）联合概率密度 $f(x,y)$；（Ⅱ）Y 的边缘概率密度 $f_Y(y)$；（Ⅱ）$P(X+Y>1)$.

【例 4】设 A、B 为事件，$P(A)=\dfrac{1}{4}$，$P(B|A)=\dfrac{1}{3}$，$P(A|B)=\dfrac{1}{2}$.

令 $X=\begin{cases}1, & A\text{发生}, \\ 0, & \overline{A}\text{发生}\end{cases}$，$Y=\begin{cases}1, & B\text{发生}, \\ 0, & \overline{B}\text{发生}\end{cases}$

求：（Ⅰ）X,Y 的联合分布律 p_{ij}；（Ⅱ）$Z=X^2+Y^2$ 的分布律.

【例 5】设 $(X,Y)\sim f(x,y)=\begin{cases}1, & 0<x<1, 0<y<2x, \\ 0, & \text{其他}\end{cases}$

求：（Ⅰ）边缘概率密度函数 $f_X(x)$，$f_Y(y)$；

（Ⅱ）$Z=2X-Y$ 的概率密度函数 $f_Z(z)$.

【例 6】设 X,Y 相互独立，$P(X=i)=\dfrac{1}{3}$，$i=-1,0,1$. $Y \sim f_Y(y)=\begin{cases}1, & 0 \le y < 1, \\ 0, & 其他\end{cases}$，

记 $Z=X+Y$，求 Z 的概率密度函数 $f_Z(z)$.

第四讲　数字特征

综述

1．求数字特征：EX 、DX 、$Cov(X,Y)$ 、ρ_{XY}

2．应用

一、概念

1．数学期望（E）与方差（D）

（1）期望定义

1）$X \sim p_i$，$EX = \sum_i x_i p_i$

2）$X \sim f(x)$，$EX = \int_{-\infty}^{+\infty} xf(x)dx$

3）$X \sim p_i$，$Y = g(X)$，$EY = \sum_i g(x_i) p_i$

4）$X \sim f(x)$，$Y = g(X)$，$EY = \int_{-\infty}^{+\infty} g(x) f(x)dx$

5）$(X,Y) \sim p_{ij}$，$Z = g(X,Y)$，$EZ = \sum_i \sum_j g(x_i, y_i) p_{ij}$

6）$(X,Y) \sim f(x,y)$，$Z = g(X,Y)$，$EZ = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f(x,y)dxdy$

（2）方差定义

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

1）定义法

1^0 $DX = \sum_i (x_i - EX)^2 p_i$

2^0 $DX = \int_{-\infty}^{+\infty} (x - EX)^2 f(x)dx$

2）公式法

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

若 EX，DX 已知，则 $EX^2 = (EX)^2 + DX$

（3）性质

1） $Ea = a$， $E(EX) = EX$

2） $E(aX + bY) = aEX + bEY$， $E\sum_{i=1}^{n} a_i X_i = \sum_{i=1}^{n} a_i EX_i$

3）若 X 、Y 相互独立，则 $EXY = EX \cdot EY$

4） $Da = 0$， $D(EX) = 0$， $D(DX) = 0$

5）若 X 、Y 相互独立，则 $D(X \pm Y) = DX + DY$

6） $D(aX + b) = a^2 DX$

7）一般地， $D(X \pm Y) = DX + DY \pm 2Cov(x, y)$

$$D(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} DX_i + 2\sum_{1 \le i \le j \le n} Cov(X_i, X_j)$$

8） $DX \le E(X - C)^2$， $\forall C$

【注】记住如下 EX， DX

①$0-1$分布　$EX = p$， $DX = p(1-p)$

② $X \sim B(n, p)$　$EX = np$， $DX = np(1-p)$

③ $X \sim P(\lambda)$　$EX = \lambda$， $DX = \lambda$

④ $X \sim Ge(p)$　$EX = \frac{1}{p}$， $DX = \frac{1-p}{p^2}$

⑤ $X \sim U[a,b]$　$EX = \frac{a+b}{2}$， $DX = \frac{(b-a)^2}{12}$

⑥ $X \sim E(\lambda)$　$EX = \frac{1}{\lambda}$, $DX = \frac{1}{\lambda^2}$

⑦ $X \sim N(\mu, \sigma^2)$　$EX = \mu$, $DX = \sigma^2$

2．协方差$Cov(X,Y)$与相关系数ρ_{XY}

（1）$Cov(X,Y)=EXY-EX\cdot EY$

1）定义法

$(X,Y)\sim p_{ij}\Rightarrow Cov(X,Y)=\sum\limits_j\sum\limits_i(x_i-EX)(y_j-EY)p_{ij}$

$(X,Y)\sim f(x,y)\Rightarrow Cov(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-EX)(y-EY)f(x,y)dxdy$

2）公式法

$Cov(X,Y)=EXY-EX\cdot EY$

①$Cov(X,X)=DX$；②$Cov(X,Y)=0\Leftrightarrow EXY=EX\cdot EY$；

③$D(X\pm Y)=DX+DY\Leftrightarrow Cov(X,Y)=0$

（2）$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{DX}\sqrt{DY}}$

$\rho=0\Leftrightarrow X,Y$不相关；$\rho\neq0\Leftrightarrow X,Y$相关.

如：设$X\sim U(-\frac{1}{2},\frac{1}{2})$，$Y=\cos X$，求$\rho_{XY}$.

【注】ρ_{XY}常称为“线性相关系数”，刻画X,Y间的线性相依程度，不刻画其“一般”函数关系.

$|\rho_{XY}|\leq1$

（3）性质

1）$Cov(X,Y)=Cov(Y,X)$

2）$Cov(aX,bY)=abCov(X,Y)$

3） $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

4） $|\rho_{xy}|\le 1$

5） $\rho_{xy}=1\Leftrightarrow P(Y=aX+b)=1(a>0)$；

$\rho_{xy}=-1\Leftrightarrow P(Y=aX+b)=1(a<0)$

考试时：$Y=aX+b$，$a>0\Rightarrow\rho_{XY}=1$；

$Y=aX+b$，$a<0\Rightarrow\rho_{XY}=-1$.

【小结】五个充要条件：

$\rho_{XY}=0\Leftrightarrow Cov(X,Y)=0$

$\Leftrightarrow EXY=EX\cdot EY$

$\Leftrightarrow D(X+Y)=DX+DY$

$\Leftrightarrow D(X-Y)=DX+DY$

若 X,Y 相互独立，则可推出上述五个结论；但反过来，不成立.

若 (X,Y) 服从二维正态分布，X,Y 相互独立$\Leftrightarrow\rho_{XY}=0$.

二、综合题解析

【例 1】设 $EX=\mu$，$\mu>0$，$\forall$ 常数 C，证明：$E(X-C)^2\ge E(X-\mu)^2=DX$.

【例 2】试验成功的概率为 $\frac{3}{4}$，失败的概率为 $\frac{1}{4}$. 独立重复试验直到成功两次为止，求试验次数 X 的期望 EX .

【例 3】T 为连续型，$t \geq 0$，$0 \leq \alpha \leq 1$，$\lambda, \mu > 0$，满足 $P(T > t) = \alpha e^{-\lambda t} + (1-\alpha)e^{-\mu t}$，求 ET， DT.

【例 4】设 X, Y 相互独立， $X \sim N(0, \frac{1}{2})$， $Y \sim N(0, \frac{1}{2})$，求 $D|X - Y|$.

【例 5】$(X,Y)\sim$

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

，则 $Cov(X^2,Y^2)=$________.

【例 6】将一枚硬币重复掷 n 次，以 X 和 Y 分别表示正面向上、反面向上的次数，则 $\rho_{XY}=$________.

【例 7】$\rho_{XY}=0.9$，$Z=X-0.4$，$\rho_{YZ}=$________.

【例 8】n 个考生的录取通知书装入 n 个信封，然后在每个信封上随意写上一个考生的姓名、地址发出. 以 S_n 表示 n 个考生中收到自己通知书的人数，求 DS_n.

【例 9】设 $X \sim f(x)$，EX，DX 存在，证明：$\forall \varepsilon > 0$，$P(|X - EX| \geq \varepsilon) \leq \frac{DX}{\varepsilon^2}$.

【例 10】设 $EX = EY = 2$，$DX = 1$，$DY = 4$，$\rho_{XY} = 0.5$，则根据切比雪夫不等式，知 $P(|X - Y| \geq 6) \leq$________.

【例 11】截至 2010 年 10 月 25 日，上海世博会参观人数超过 7000 万人.游园最大的痛苦就是人太多.设游客到达中国馆有三条路径.
沿第一条路走 3 小时到达；
沿第二条路走 5 小时又回到原处；
沿第三条路走 7 小时又回到原处.
游客等可能选其一种路径，求游客平均要用多少时间到达中国馆？

第五讲　大数定律与中心极限定理

综述

1．依概率收敛

2．三大定律，两大定理

一、依概率收敛

设$\{X_n\}$为一随机变量序列，X为一随机变量（或a为常数）.

若$\forall \varepsilon > 0$，恒有$\lim\limits_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$或$\lim\limits_{n\to\infty} P(|X_n - a| < \varepsilon) = 1$，则称$\{X_n\}$依概率收敛于$X$（或$a$），记$X_n \xrightarrow{P} X$（或$a$）.

【例】设$\{X_n\}$，$X_n \sim f_n(x) = \dfrac{n}{\pi(1+n^2x^2)}$，$-\infty < x < +\infty$，证明：$X_n \xrightarrow{P} 0$.

二、大数定律

1．切比雪夫大数定律

设$\{X_n\}$是相互独立的随机变量序列，若方差DX_k存在且一致有上界，则

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} EX_i$$

2．伯努利大数定律

设u_n是n重伯努利试验中事件A发生的次数，在每次试验中A发生的概率为p，则

$$\frac{u_n}{n}\xrightarrow{P}p$$

3．辛钦大数定律

设$\{X_n\}$是独立同分布的随机变量序列，若$EX_n=\mu$存在，则$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\mu$.

【例 1】设$X_1,X_2,\cdots,X_n,\cdots$是相互独立的序列，$X_n\sim E_X(n)$，则下列序列中不服从切比雪夫大数定律的是（　）

（A）$X_1,\frac{1}{2}X_2,\cdots,\frac{1}{n}X_n,\cdots$　　（B）$X_1,X_2,\cdots,X_n,\cdots$

（C）$X_1,2X_2,\cdots,nX_n,\cdots$　　（D）$X_1,2^2X_2,\cdots,n^2X_n,\cdots$

【例 2】设序列$X_1,X_2,\cdots,X_n,\cdots$相互独立，根据辛钦大数定律，当$n\to\infty$时，$\frac{1}{n}\sum_{i=1}^{n}X_i$依概率收敛于其数学期望，只要$\{X_n\}$（　）

（A）有相同的数学期望　　（B）服从同一离散型分布

（C）服从同一泊松分布　　（D）服从同一连续型分布

【例 3】设$X\sim E_X(2)$，$X_1,X_2,\cdots,X_n$独立同分布于$E_X(2)$，则当$n\to\infty$时，

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} \underline{\qquad\qquad}.$$

三、中心极限定理（$n \to \infty$）

【定理】(列维——林德伯格中心极限定理，即独立同分布中心极限定理)假设$\{X_n, n \geq 1\}$是独立同分布的随机变量序列，如果

$EX_n = \mu$，$DX_n = \sigma^2 > 0$（$n \geq 1$）存在，则$\{X_n, n \geq 1\}$服从中心极限定理，即对任意的实数x，有

$$\lim_{n\to\infty} P\{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leq x\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}x} dt = \Phi(x).$$

【定理】(棣莫弗——拉普拉斯中心极限定理，即二项分布以正态分布为其极限分布定理)假设随机变量$Y_n \sim B(n, p)$，（$0 < p < 1$，$n \geq 1$），则对任意实数x，有

$$\lim_{n\to\infty} P\{\frac{Y_n - np}{\sqrt{np(1-p)}} \leq x\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-x/2} dt = \Phi(x).$$

【例】设$X_1, X_2, \cdots, X_n$为独立同分布的随机变量序列，且均服从参数为λ（$\lambda > 1$）的指数分布. 记$\Phi(x)$为标准正态分布函数，则（　）

（A）$\lim_{n\to\infty} P(\frac{\sum_{i=1}^{n} X_i - n\lambda}{\sqrt{n}\lambda} \leq x) = \Phi(x)$　　（B）$\lim_{n\to\infty} P(\frac{\sum_{i=1}^{n} X_i - n\lambda}{\sqrt{n\lambda}} \leq x) = \Phi(x)$

（C）$\lim\limits_{n\to\infty}P(\dfrac{\lambda\sum\limits_{i=1}^{n}X_i-n}{\sqrt{n}}\le x)=\Phi(x)$ （D）$\lim\limits_{n\to\infty}P(\dfrac{\sum\limits_{i=1}^{n}X_i-\lambda}{\sqrt{n\lambda}}\le x)=\Phi(x)$

第六讲　数理统计

综述

1．统计量的分布（正态分布、χ^2 分布、t 分布、F 分布）

2．求统计量的数字特征

3．估计与评价

一、总体与样本

1．总体

研究对象的全体称为总体，组成总体的每一个元素称为个体．在对总体进行统计研究时，我们所关心的是表征总体状况的某个(或某几个)数量指标 X (可以是向量)和该指标在总体中的分布情况.我们把总体与数量指标 X 可能取值的全体所组成的集合等同起来，或直接将总体与随机变量 X 等同起来，说“总体 X ”. 所谓总体的分布就是指随机变量 X 的分布.

2．样本——简单随机样本　（独立同分布）

n 个相互独立且与总体 X 具有相同概率分布的随机变量 $X_1, X_2, \cdots, X_n$ 的整体 $(X_1, X_2, \cdots, X_n)$ 称为来自总体 X ，容量为 n 的一个简单随机样本，简称为样本.

二、统计量

设 $(X_1, \cdots, X_n)$ 为总体 X 的样本，$g(x_1, \cdots, x)$ 为 n 元函数，如果 g 中不含任何未知参数，则称 $g(X_1, \cdots, X_n)$ 为样本 $(X_1, \cdots, X_n)$ 的一个统计量．若 $(x_1, \cdots, x_n)$ 为样本值，则称 $g(x_1, \cdots, x_n)$ 为 $g(X_1, \cdots, X_n)$ 的观测值.

常用统计量

1、样本均值 $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$

2、样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n-1}[\sum_{i=1}^{n}X_i^2-n\overline{X}^2]$

样本标准差 $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$

三、四大分布

1．正态分布 $X\sim N(\mu,\sigma^2)$

$EX=\mu$，$DX=\sigma^2$

标准化 $\frac{X-\mu}{\sigma^2}\sim N(0,1)$

若 $P(u\ge u_\alpha)=\alpha$，称 u_α 为上 α 分位数.（$u=\frac{X-\mu}{\sigma^2}$）

2． χ^2 分布

若随机变量 $(X_1,\cdots,X_n)$ 独立同分布于标准正态分布，则随机变量 $X=\sum_{i=1}^{n}X_i^2$ 服从自由度为 n 的 χ^2 分布，记为 $X\sim\chi^2(n)$.

若 $P(X\ge\chi_\alpha^2(n))=\alpha$，则称 $\chi_\alpha^2(n)$ 为上 α 分位数.

$EX=n$，$DX=2n$

3． t 分布

若 $X\sim N(0,1)$，$Y\sim\chi^2(n)$，且 X，Y 相互独立，则 $t=\frac{X}{\sqrt{Y/n}}\sim t(n)$.

若 $P(t\ge t_\alpha(n))=\alpha$，则称 $t_\alpha(n)$ 为上 α 分位数.

4． F 分布

若 $X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，X，Y 独立，则 $F=\frac{X/n_1}{Y/n_2}\sim F(n_1,n_2)$.

$\frac{1}{F}=\frac{Y/n_2}{X/n_1}\sim F(n_2,n_1)$

四、正态下的常用分布

设 $X_1,\cdots,X_n$ 独立同分布于正态分布 $N(\mu,\sigma^2)$，

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i，\quad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

① $\overline{X}$ 与 S^2 独立，且 $E\overline{X}=\mu$， $D\overline{X}=\frac{1}{n}\sigma^2$， $ES^2=\sigma^2$.

② $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$ （σ^2 已知）

③ $\dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ （σ^2 未知）

【例】设一批零件的长度服从正态分布 $N(\mu,\sigma^2)$，其中 μ,σ^2 均未知，现随机抽取 16 个零件，测得 $\overline{x}$=20（cm）， s=1（cm），则 μ 的置信度为 0.90 的置信区间为________.

五、单独讨论四大分布（N，χ^2，t，F）的构成

①$\{X_i\}$独立同分布于$N(0,1)\Rightarrow\sum\limits_{i=1}^{n}X_i^2\sim\chi^2(n)$

②$X\sim N(0,1)$，$Y\sim\chi^2(n)$，X，Y相互独立$\Rightarrow\dfrac{X}{\sqrt{Y/n}}\sim t(n)$

③$X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，X，Y相互独立$\Rightarrow\dfrac{X/n_1}{Y/n_2}\sim F(n_1,n_2)$

【例】设X，Y相互独立且服从于正态分布$N(0,3^2)$，$X_1,\cdots,X_9$和$Y_1,\cdots,Y_9$分别是来自总体X和Y的简单随机样本，则统计量$U=\dfrac{X_1+\cdots+X_9}{\sqrt{Y_1^2+\cdots+Y_9^2}}\sim$________.

六、统计量的数字特征

$Y=g(X_1,X_2,\cdots,X_n)$

【例】设$X_1,X_2,\cdots,X_n(n>2)$来自标准正态分布总体的简单随机样本. 记$\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$，$Y_i=X_i-\overline{X}$，$i=1,2,\cdots,n$.

求（Ⅰ）DY_i，$i=1,2,\cdots,n$；（Ⅱ）$Cov(Y_1,Y_n)$；（Ⅲ）$P(Y_1+Y_n\le 0)$.

七、点估计与评价标准

1．矩估计

①对于一个参数，令 $\overline{X}=EX$，即 $\frac{1}{n}\sum_{i=1}^{n}X_i=\begin{cases}\sum_i x_i p_i & X_i\sim p_i\\ \int_{-\infty}^{+\infty}xf(x)dx & X_i\sim f(x)\end{cases}$

②对于两个参数：

1^0 令 $\overline{X}=EX$；

2^0 令 $\frac{1}{n}\sum_{i=1}^{n}X_i^2=EX^2$ 或 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=DX$

【例 1】设 $X\sim f(x,\theta)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1\\ 0, & \text{其他}\end{cases}$，求 θ 的矩估计量.

【例 2】设总体 $X \sim f(x,\theta,\lambda)=\begin{cases}\dfrac{1}{\lambda}e^{-\frac{1}{\lambda}(x-\theta)}, & x\geq\theta, \\ 0, & x<\theta\end{cases}$，$\lambda>0$，由样本 $X_1,X_2,\cdots,X_n$，求 λ， θ 的矩估计量.

2．最大似然估计

引例

求解步骤：

（1）写出似然函数：$L(x_1,x_2,\cdots,x_n,\theta)=\begin{cases}\prod\limits_{i=1}^{n}p(x_i,\theta)\\ \prod\limits_{i=1}^{n}f(x_i,\theta)\end{cases}$

（2）令 $\dfrac{dL}{d\theta}=0$ 或 $\dfrac{d\ln L}{d\theta}=0\Rightarrow\hat{\theta}$

当 $L(\theta)$ 单调时，用定义.

【注】若为 $L(x_1,x_2,\cdots,x_n,\theta_1,\theta_2)$，

令 $\dfrac{\partial L}{\partial\theta_1}=0$，$\dfrac{\partial L}{\partial\theta_2}=0$（或 $\dfrac{\partial\ln L}{\partial\theta_1}=0$，$\dfrac{\partial\ln L}{\partial\theta_2}=0$）

【例 1】设 X 的分布律为 $X\sim\begin{pmatrix}0 & 1 & 2 & 3\\ \theta^2 & 2\theta(1\text{-}\theta) & \theta^2 & 1-2\theta\end{pmatrix}$，其中 $\theta(0<\theta<\dfrac{1}{2})$ 为未知参数. 利用总体 X 的如下样本值：3，1，3，0，3，1，2，3，求 θ 的矩估计值和最大似然估计值.

【例 2】设 $X \sim F(x,\alpha,\beta)=\begin{cases}1-(\dfrac{\alpha}{x})^{\beta}, & x>\alpha\\ 0, & x\le\alpha\end{cases}$，其中参数 $\alpha>0$，$\beta>1$. 设

$X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本.

（Ⅰ）当 $\alpha=1$ 时，求 β 的矩估计量；

（Ⅱ）当 $\alpha=1$ 时，求 β 的最大似然估计量；

（Ⅲ）当 $\beta=2$ 时，求 α 的最大似然估计量.

【例 3】设 $X \sim U[0,\theta]$，$\theta>0$ 为未知参数，求 θ 的最大似然估计量.

3．估计量的评价标准

【引例】设 $X \sim U[0,\theta]$，求 θ 的矩估计量、最大似然估计量.

①无偏性

若参数 θ 的估计量 $\hat{\theta}$ 对一切 n，n 有 $E(\hat{\theta})=\theta,$ 则称 $\hat{\theta}$ 为 θ 的无偏估计量，否则称为有偏估计量.

②有效性

若 $\hat{\theta}_1$ 与 $\hat{\theta}_2$ 均为 θ 的无偏估计量，

当 $D\hat{\theta}_1 < D\hat{\theta}_2$ 时，$\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

③一致性（相合性）（只针对大样本 $n \to +\infty$）

设 $\hat{\theta}$ 为 θ 的估计量，如果对任意 $\varepsilon > 0$，有 $\lim\limits_{n\to\infty} P(\left|\hat{\theta}-\theta\right| \ge \varepsilon)=0$ 或 $\lim\limits_{n\to\infty} P(\left|\hat{\theta}-\theta\right| < \varepsilon)=1$ 即 $\hat{\theta} \xrightarrow{P} \theta(n\to\infty)$，则称 $\hat{\theta}$ 为 θ 的一致估计量(或相合估计量).

【例 1】设总体 $X\sim f(x,\theta)=\begin{cases}2e^{-2(x-\theta)}, & x>\theta\\ 0, & x\le\theta\end{cases}$，其中 $\theta>0$，为未知参数，从总体中取简单随机样本 $X_1,X_2,\cdots,X_n$，记 $\hat{\theta}=\min\{X_1,X_2,\cdots,X_n\}(=X_{(1)})$.

（Ⅰ）求 $X\sim F(x)$；

（Ⅱ）求 $\hat{\theta}\sim F_{\hat{\theta}}(x)$；

（Ⅲ）若用 $\hat{\theta}$ 作为 θ 的估计量，讨论它是否具有无偏性.

【例 2】设 $X\sim F(x;\theta)=\begin{cases}1-e^{-\frac{x^2}{\theta}}, & x\ge 0,\\ 0, & x<0\end{cases}$，其中 θ 是未知参数，$\theta>0$. $X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本.

（Ⅰ）求 EX 与 EX^2；

（Ⅱ）求 θ 的最大似然估计量 $\hat{\theta}$；

（III）是否存在实数a，使得对$\forall\varepsilon>0$，都有$\lim\limits_{n\to\infty}P(\left|\hat{\theta}-a\right|\geq\varepsilon)=0$？

八、假设检验

1．回顾

$$P(\left|\overline{X}-\mu\right|<u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})=1-\alpha \qquad （\sigma^2\text{已知}）$$

$$P(\left|\overline{X}-\mu\right|<t_{\frac{\alpha}{2}}(n)\cdot\frac{S}{\sqrt{n}})=1-\alpha \qquad （\sigma^2\text{未知}）$$

2．已知说法：H_0：$\mu=\mu_0$，H_1：$\mu\neq\mu_0$

【注】若$P(A)=\alpha$（显著性水平，人为取$\alpha=0.1$，$0.05, 0.01$，…）称A为小概率事件，我们认为，小概率事件在一次试验中是不会发生的. 若发生了，则拒绝原假设.

①设总体$X\sim N(\mu,\sigma^2)$，σ^2已知，μ未知

H_0： $\mu=\mu_0$， H_1： $\mu\neq\mu_0$，构造一个水平为α 的检验

②σ^2 未知，其他同①，构造一个水平为α 的检验

③总体 $X\sim N(\mu,\sigma^2)$，σ^2 已知， μ 未知

1^0 H_0： $\mu\leq\mu_0$， H_1： $\mu>\mu_0$，构造一个水平为α 的检验

2^0 H_0： $\mu \geq \mu_0$， H_1： $\mu < \mu_0$，其他同③，构造一个水平为α 的检验

④$1^0$若σ^2未知，其他同③$1^0$

H_0： $\mu \leq \mu_0$， H_1： $\mu > \mu_0$

2^0若σ^2未知，其他同③$2^0$

H_0： $\mu \geq \mu_0$， H_1： $\mu < \mu_0$

【例】已知某机器生产出的零件长度X（cm）服从正态分布$N(\mu,\sigma^2)$，μ，σ^2均未知，

现从中取容量 $n=16$ 的一个样本，测得 $\overline{x}=10$ （cm）， $s^2=0$ ， $t_{0.025}(15)=2.132$.

（Ⅰ）求均值 μ 的置信度为 0.95 的置信区间；

（Ⅱ）在 $\alpha=0.05$ 下，检验 H_0： $\mu=9.7$ ， H_1： $\mu\neq 9.7$.