写在前面

> 本文为南方科技大学MA212概率论与数理统计课程2022秋季学期前半学期的知识点整理，大部分内容摘选自MA212-03班的课件，编者仅进行了一定转述和搬运工作。由于复习材料的目的性强，本文对考点有一定的强调和浓缩，忽略了一些非考点，但也并不能视为与考纲完全相符，同时可能存在一定错误。阅读本文并不能让你两三天速通概统，本文的定位在于用一篇文章看完课件内容，因此建议读者在学完该部分内容后借本文复习概念，并结合题目练习。
>
> 祝你在这门课程取得优秀成绩！
>
> 计算机科学与技术21级 樊斯特(12111624)
>
> 2023.1

由于编者能力有限，恳请各位读者在阅读过程中勘误，如果您对其中内容有疑问或发现错误，抑或是希望联系编者，欢迎通过该QQ联系：3425811925，非常感谢！

Chapter 1

§1.2 样本空间

试验

随机试验(试验)：

- 可以在相同的条件下重复进行
- 试验前知道所有可能结果
- 试验前无法确定会出现哪个结果

样本空间Ω：试验的全部样本点构成的集合

样本点Ω：不可分的试验结果

事件

随机事件(事件)：满足一定条件的样本点的集合

基本事件：一个样本点构成的单点集

必然事件：每次试验都总发生的事件

不可能事件：每次试验都不会发生的事件$A = \{A|A \subset \boldsymbol{\Omega}, A\text{是事件}\}$

事件运算：

$A \subset B$	$A \cup B$ **和**	$A \cap B$ **积**	$A - B$ **差**
A发生必导致B发生	A、B 至少有一个发生	A、B同时发生	A发生B不发生

- 若$A \supset B$，则$A-B$为**真差**
- 若$A \cap B = \phi$，则AB**互不相容(互斥)**
- 若$A \cup B = \mathbf{\Omega}$且$A \cap B = \phi$，则$AB$互为**逆事件/对立事件**，记作$A = B^C = \bar{B}$

事件运算定律：

交换、结合、分配、德摩根

可列：无穷集S可与自然数集建立双射，可表示为$S = \{s1, s2, \ldots\}$，“最小的无穷”

至多可列：可列/有限

§1.3 概率测度

频率

设 A 为一随机事件 ,在相同条件下进行 n 次重复试验

频数：$n_A = n$次试验中A发生的次数

频率：$f_{\mathbf{N}}(A) = \frac{n_A}{n}$

频率的基本性质：

- $0 \leq f_{\mathbf{N}}(A) \leq 1$
- $f_{\mathbf{N}}(\mathbf{\Omega}) = 1$
- 若$A_1, A_2, \ldots A_m$两两不相容，则

  $f_{\mathbf{N}}(\bigcup\limits_{i=1}^{m} A_i) = \sum\limits_{i=1}^{m} f_{\mathbf{N}}(A_i)$

概率

概率：事件A发生的可能性大小$P(A) = \lim\limits_{n \to +\infty} f_{\mathbf{N}}(A)$

概率空间：样本空间+事件域+概率，$\{\mathbf{\Omega}，\mathcal{A}，P\}$

概率的公理化定义：

- **非负性**：$0 \leq P(A) \leq 1，(\forall A \in \mathcal{A})$
- **规范性**：$P(\mathbf{\Omega}) = 1$
- **可列可加性**：对于两两不相容的事件列$\{A_k\}_{k=1}^{\infty}$有

  $P(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$

概率的基本性质：

- $P(\emptyset) = 0$
- **有限可加性**：对于两两不相容的事件列$\{A_k\}_{k=1}^{n}$有

  $P(\bigcup\limits_{i=1}^{n} A_i) = \sum\limits_{i=1}^{n} P(A_i)$
- 若$A \subset B$，则$P(B-A) = P(B) - P(A)$，$P(B) \geq P(A)$
- $0 \leq P(A) \leq 1$
- $P(\bar{A}) = 1 - P(A)$
- **多事件的加法定律**

§1.4 概率计算：计数方法

古典概型(等可能概型)：仅有有限个等可能出现的基本结果的随机试验

有利场合：导致事件发生的方式个数

排列与组合：$A_n^k=\frac{n!}{(n-k)!}$，$\binom{n}{k}=C_n^k=\frac{n!}{k!(n-k)!}$，$\binom{n}{n_1\ n_2\ \dots\ n_r}=\frac{n!}{n_1!n_2!\dots n_r!}$

乘法原理&加法原理：略

几何概型：e.g. 向平面有界区域$\mathbf{\Omega}$投掷一个点，点落在可测量的平面区域A的试验

§1.5 条件概率

条件概率基础

Monty Hall Problem(三门问题)：

条件概率：A、B是两个事件，且$P(B)>0$，记$P(A|B)=\frac{P(AB)}{P(B)}$为在事件$B$发生的条件下事件$A$发生的条件概率

- $P(B)=0$时条件概率无意义
- B发生带来的“信息”对A的“推断”的新认识
- $A|B$不是一个事件
- $P(A|B)\geq P(AB)$

条件概率的基本性质：

- **非负性**：$P(A|B)\geq 0$
- **规范性**：$P(\mathbf{\Omega}|B)=1$
- **可列可加性**：对于两两不相容的事件列$\{A_k\}$有
  $P(\bigcup_{i=1}^{\infty}A_i|B)=\sum_{i=1}^{\infty}P(A_i|B)$

样本空间$\widetilde{\mathbf{\Omega}}=B\cap\mathbf{\Omega}$，事件域$\widetilde{\mathcal{A}}=\{B\cap A|A\in\mathcal{A}\}$，概率$P_B(A)=P(A|B)$

条件概率相关公式

乘法定律(公式)：$P(AB)=P(A|B)\cdot P(B)$

分划：事件列$\{B_k\}$两两不相容，且$\bigcup_{i=1}^{n}B_i=\mathbf{\Omega}$

全概率公式：B_k是$\mathbf{\Omega}$的一个分划，对任意事件A有

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

贝叶斯公式：通过原因概率和先验概率计算后验概率

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

§1.6 独立性

相互独立(独立)：两两独立，三三独立...

$P(A_{i_1}A_{i_2}\dots A_{i_k})=P(A_{i_1})P(A_{i_2})\dots P(A_{i_k})$

$(1\leq i_1<\dots<i_k\leq n，k=2\dots n)$

- AB**独立**与AB**不相容**不能同时成立

$P(AB)=P(A)P(B)\neq 0$

- AB独立，则$A\bar{B}$独立，$\bar{A}B$独立，$\bar{A}\bar{B}$独立
  $P(AB)=P(A)(1-P(\bar{B}))=P(A)-P(A)P(\bar{B})$
  $P(A)P(\bar{B})=P(A)-P(AB)=P(A-AB)=P(A\bar{B})$
- 两两独立是相互独立的必要不充分条件

Chapter 2

§2.1 离散随机变量

离散型随机变量基础

随机变量：令Ω为一个样本空间，令X是定义在Ω上的一个实函数，则称X为一个(一维)随机变量。随机事件可以用随机变量的取值来表示。

离散型随机变量：随机变量X仅取至多可列个值

概率质量函数/频率函数(PMF)：所有可能的取值→取各个值的概率

$P\{X=x_k\}=p(x_k)\ (k=1,2,\dots)$

频率函数的本质特征(充分必要)：

- $p(x_k\geq 0)$，$k=1,2..$
- $\sum_{k=1}^{\infty}p(x_k)=1$

累积分布函数/分布函数(CDF)：$F(x)=P\{X\leq x\},-\infty<x<\infty$

分布函数的本质特征(充分必要)：

- $F(x)$是单调不减函数
- $0\leq F(X)\leq 1$且$F(-\infty)=0,F(\infty)=1$
- $F(x)$右连续，即$\lim_{t\to x^+}F(t)=F(x)$

分布函数的其他性质：

- $P\{a<X\leq b\}=F(b)-F(a),\ a<b$
- $P\{X=c\}=\lim_{\Delta t\to 0^+}[F(c)-F(c-\Delta t)]=F(c)-F(c-0)$

离散随机变量的分布

单点分布(退化分布)：事件几乎处处发生，$P(X=c)=1$，记为$X\overset{a.e}{=}c$或$X=c\ (a.e)$

(0-1)两点分布(伯努利随机变量)：$P\{X=1\}=p$，$P\{X=0\}=1-p$

伯努利试验：只产生两个结果$A,\bar{A}$的试验

二项分布：$P\{X=k\}=C_n^kp^k(1-p)^{n-k}\ (k=0,1,2,\dots n)$，记为$X\sim b(n,p)$

- $b(k;n,p):=C_n^kp^k(1-p)^{n-k}$
- 当$k<(n+1)p$时，$b(k;n,p)$随k增加而增加
- 当最可能出现次数$k=(n+1)p$为正整数时，中心项$b(k;n,p)=b(k-1;n,p)$

几何分布：前面的$k-1$次伯努利试验失败，第k次试验成功，故

$p(k)=P\{X=k\}=(1-p)^{k-1}p,\ k=1,2,3,\dots$

负二项分布：连续独立地试验直到成功r次为止，停止时做了k次试验

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

超几何分布：n个球，r黑，$n-r$白，盒中不放回地抽取m球，X代表抽到的黑球个数

$P\{X=k\}=\frac{C_r^kC_{n-r}^{m-k}}{C_n^m},\ k=1,...m$

泊松分布：参数$\lambda>0$，r.v. X的取值为0,1,2...，取值概率为

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

记为$X\sim\pi(\lambda)$或$X\sim P.(\lambda)$

泊松分布的性质：

- $P\{X=k\}>0,\ k=0,1,2,...$
- $\sum_{k=0}^{\infty}P\{X=k\}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{\lambda}\cdot e^{-\lambda}=1$

泊松流与泊松分布：记时间间隔$(0,t]$中出现的质点数为X，则$X\sim P(\lambda t)$，其中泊松强度$\lambda>0$

泊松定理：设$\lambda>0$为一常数，n为正整数，$\lim_{n\to\infty}np_n=\lambda$，则对任一非负整数$k$有

$\lim_{n\to\infty}C_n^kp_n^k(1-p_n)^{n-k}=\frac{\lambda^ke^{-\lambda}}{k!}$

当n很大p很小时，$C_n^kp^k(1-p)^{n-k}\approx\frac{\lambda^ke^{-\lambda}}{k!}$

§2.2 连续随机变量

连续型随机变量基础

连续型随机变量：若r.v. X的分布函数能够表示为

$F(x)=\int_{-\infty}^{x}f(t)dt,\ -\infty<x<\infty$

其中$f(t)\geq0$且可积，则称X为连续性r.v.

概率密度函数(PDF)：$f(t)$，反映了概率集 中在该点附近的程度，不代表概率

密度函数的本质特征：

- $f(t)\geq0$
- $\int_{-\infty}^{\infty}f(t)dt=1$

密度函数的性质：

- $\forall x_1<x_2$有

  $P\left\{x_1<X\leq x_2\right\}=F\left(x_2\right)-F\left(x_1\right)=\int_{x_1}^{x_2}f(x)dx$
- 在$f(x)$的连续点处有$f(x)=F'(x)$
- 补充几个不可导处点值不影响积分值
- $\forall c$，$P\{X=c\}=0$

p分位数：对于$X\sim f(x)$，若$\forall0<p<1$，存在常数x_p满足$F(x_p)=p$，则x_p为密度函数$f(x)$的p分位数

概率函数：

$$f(x)\begin{cases}\text{对离散型 } r.v\ X \text{ 表示频率函数,即}\\ f(x)=P\{X=x\},x=x_1,x_2,\cdots\\ \text{对连续型 } r.v\ X \text{ 表示密度函数,即有}\\ f(x)\geq 0,\int_{-\infty}^{\infty}f(x)dx=1\end{cases}$$

连续随机变量的分布

正态分布

正态分布 $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，其中$-\infty<\mu<\infty$，$\sigma>0$，记为$X\sim\mathbf{N}(\mu,\sigma^2)$

正态分布函数的性质：

- $f(\mu+x)=f(\mu-x)$，即$f(x)$关于$x=\mu$对称
- $x<\mu$，$f'(x)>0$，$f(x)\uparrow$
  $x>\mu$，$f'(x)<0$，$f(x)\downarrow$
  极大值$f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$
- μ不改变形状，改变对称轴位置；σ不改变对称轴，改变图像形状

标准正态分布：$X\sim\mathbf{N}(0,1)$

$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$， $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$

若$X\sim\mathbf{N}(\mu,\sigma^2)$，则$Z=\frac{X-\mu}{\sigma}\sim\mathbf{N}(0,1)$

均匀分布

均匀分布：若r.v. X的密度函数为

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

则称X服从区间(a,b)上的均匀分布，记为$X\sim\mathbf{U}(a,b)$

- $\forall(c,c+L)\subset(a,b)$，$P\{c<X\leq c+L\}=\int_{c}^{c+L}\frac{dx}{b-a}=\frac{L}{b-a}$

指数分布

指数分布：若r.v. X的密度函数为

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

则称X服从参数为$\lambda>0$的指数分布，记为$X\sim\mathbf{EXP}(\lambda)$

- X的分布函数为$F(x)=\begin{cases}1-e^{-\lambda x}, & x>0\\ 0, & x\leq 0\end{cases}$
- 泊松流中第一个质点出现的时间服从指数分布：$P\{Y>t\}=P\{X=0\}=e^{-\lambda t}$
- λ称为失效率，λ^{-1}表示平均寿命
- 无记忆性：$P\{X>s+t|X>s\}=P\{X>t\}$

伽马分布、贝塔分布(了解)

伽马分布：一般地，设X为连续型r.v，概率密度为

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

其中$r>0,\lambda>0$为常数，则称X服从参数为(r,λ)的Γ分布，记为$X\sim\Gamma(r,\lambda)$，此处
$\Gamma(r)=\int_0^\infty x^{r-1}e^{-x}dx \quad (r>0)$

当r为自然数时，$\Gamma(r)=(r-1)!$

伽马分布的应用：假如在$(0,t]$内元件受到的冲击次数N_t是一个Poisson流。 当r是任意自然数时，元件寿命就是第r次冲击来到的时间， 记为X，则$X\sim\Gamma(r,\lambda)$，它的可靠度函数

$R(t)=P\{X>t\}=\sum_{i=0}^{r-1}P\{N_t=i\}=e^{-\lambda t}\sum_{i=0}^{r-1}\frac{(\lambda t)^i}{i!}$

贝塔分布：

贝塔密度用来刻画$[0,1]$区间上的随机变量：

$f(u)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}u^{a-1}(1-u)^{b-1},\quad 0\le u\le 1$

特别地，$a=b=1$时为均匀分布

§2.3 随机变量的函数

随机变量的函数：对于r.v. X，$g(\cdot):\boldsymbol{R}\to\boldsymbol{R}$是一个函数，将$X$代入函数$g(\cdot)$，得到新的随机变量$\boldsymbol{Y}=g(X)$： $\boldsymbol{\Omega}\to\boldsymbol{R}$，即$\boldsymbol{Y}(\omega)=g(X(\omega)),\forall\omega\in\boldsymbol{\Omega}$

随机变量$\boldsymbol{Y}$称为随机变量X的函数

离散型 r.v. 函数的频率函数

1. 列出Y的可能取值
2. 找出$\{Y=y\}$的等价事件$\{X\in D\}$
3. $P\{Y=y\}=P\{X\in D\}$

离散→离散：

$y=g(x_1)=g(x_2)=\ldots=g(x_k)$，$P\{\boldsymbol{Y}=y\}=\sum_{i=1}^k P\{X=x_i\}$

连续→离散：

$Y=c,a<X\le b$，$P\{Y=c\}=P\{a<X\le b\}=F_X(b)-F_X(a)$

连续型 r.v. 的频率函数

1. 求Y的分布函数$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}$
2. 转化为关于r.v. X的概率计算问题(用到$y=g(x)$的性质)
3. 求导得$f_Y(y)=\frac{d}{dy}F_Y(y)$

定理：设r.v. X的密度函数为$f(x)$，又$y=g(x)$是严格单调函数，其反函数$h(y)=g^{-1}(y)$连续可导，则$Y=g(X)$的密度函数为：

$$f_Y(y)=\begin{cases}|h'(y)|\cdot f(h(y)), & h(y)\text{ 有意义}\\ 0, & \text{其他}\end{cases}$$

结论：正态r.v.的线性函数仍是正态r.v.

推广的定理：设r.vX的密度函数为$f(\boldsymbol{x})$,又函数$g(x)$在若干互不相交的区间 (a_i, b_i)上逐段严格单调, 且其反函数$h_i(y)$均连续可导,则$Y=g(X)$的密度函数为

$$f_Y(y)=\begin{cases}\sum_{i=1}\left|h_i'(y)\right|\cdot f\left(h_i(y)\right), & h_1(y), h_2(y), \cdots \text{ 有意义} \\ 0, & \text{其他}\end{cases}$$

均匀分布与其他连续分布的关系：

设r.v. X的密度为$f(x)$, 分布函数为$F(x)$。其中$F(x)$在某区间I上严格递增,I的左端点处 $F=0$, 右端点处$F=1$。I 可以是有界区间, 也可以是无界区间. 因此, $F^{-1}(x)$在I上都有定义.

1. 令$Z=F(X)$，则$Z\sim\mathbf{U}(0,1)$。
2. 令$U\sim\mathbf{U}(0,1)$，$X=F(U)$，那么X的分布函数是$F^{-1}(x)$。

应用：要生成分布函数为$F(x)$的r.v., 只需将F^{-1}作用在均匀分布的随机数上即可.

Chapter 3

§3.1 联合累积分布函数

二维随机变量(向量)：设$\mathbf{\Omega}$为样本空间,$X=X(\omega), Y=Y(\omega)$ $(\omega\in\mathbf{\Omega})$ 是 定义在$\mathbf{\Omega}$上的两个随机变量，记

$(X,Y)\triangleq(X(\omega),Y(\omega))(\omega\in\boldsymbol{\Omega})$

称(X,Y)为二维随机变量（向量）。

联合累积分布函数：设(X,Y)是二维r.v.，$\forall x,y\in(-\infty,\infty)$，定义

$F(x,y)\triangleq P(\{X\leq x\}\cap\{Y\leq y\})=P\{X\leq x,Y\leq y\}$

则称$F(x,y)$为二维r.v. (X,Y)的联合累积分布函数。

联合累积分布函数的本质特征：

1. $\forall x_0$，$F(x_0,y)$是y的单调不减函数

   $\forall y_0$，$F(x,y_0)$是x的单调不减函数
2. $0\leq F(x,y)\leq 1$，且

   $F(\infty,\infty)=1$，$F(-\infty,-\infty)=0$

   $\forall x,y$，$F(-\infty,y)=F(x,-\infty)=0$
3. $F(x,y)$关于x和y分别右连续
4. $\forall x_1<x_2, y_1<y_2$有

   $F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)\geq 0$

   $=P\{x_1<X\leq x_2, y_1<Y\leq y_2\}$

n维随机变量及其联合分布函数：Trivial.

§3.2 二维离散随机变量

联合频率函数：$P\{X=x_i,Y=y_j\}=p(x_i,y_j)\triangleq p_{ij}$

边际分布函数：$F_X(x)=P\{X\leq x\}=P\{X\leq x,Y\leq\infty\}=F(x,\infty)$

随机变量的边际分布完全由其联合分布确定

边际频率函数：

$P\{X=x_i\}=\sum_{j=1}^{\infty} p_{ij} \triangleq p_{i\cdot}$

$P\{Y=y_j\}=\sum_{i=1}^{\infty} p_{ij} \triangleq p_{\cdot j}$

n维离散随机变量的频率函数：trivial.

多项分布：假设进行n次独立试验, 每次试验有r种可能的结果, 各自出现的概率分别为$p_1,p_2,\ldots p_r$。

令N_i是第n次试验出现第i种试验结果的所有次数，其中$i=1..\,.r$。$N_1,N_2,\ldots N_r$的联合频率函数是

$p\left(n_1,\ldots,n_r\right)=\binom{n}{n_1,\cdots,n_r} p_1^{n_1} p_2^{n_2}\cdots p_r^{n_r}$

$N_i\sim b(n,p_i)$，$p_{N_i}\left(n_i\right)=\binom{n}{n_i} p_i^{n_i}(1-p_i)^{n-n_i}$

§3.3 二维连续随机变量

二维连续型r.v.

二维连续型随机变量：设r.v. (X,Y)的联合分布函数为$F(x,y)=P\{X\leq x,Y\leq y\}$，若存在非负可积函数$f(x,y)\geq 0$使得

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

则称(X,Y)为二维连续型随机变量

联合概率密度：上述$f(x,y)$称为X,Y的联合概率密度

密度函数的本质特征：

1. $f(x,y)\geq 0\ (\forall(x,y)\in R^2)$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u,v)dudv=1$

密度函数的基本性质：

1. $\forall D\subset R^2$，其中D为由逐段光滑曲线围成的平面区域

   $P\{(X,Y)\in D\}=\iint\limits_{D} f(x,y)dxdy$
2. 在$f(x,y)$的连续点处，有

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

约定：

$(X_1,X_2,\cdots,X_n)\sim F\left(x_1,x_2,\cdots,x_n\right)$，$n$维随机变量的分布函数是$F$

$(X_1,X_2,\cdots,X_n)\sim f\left(x_1,x_2,\cdots,x_n\right)$，$n$维随机变量的频率函数是$f$

边际分布函数：

$F_X(x)=P\{X\leq x\}=P\{X\leq x,Y\leq\infty\}=\int_{-\infty}^{x}\int_{-\infty}^{\infty} f(u,y)dydu$

边际密度函数：

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

n维边际密度函数：

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

二维联合分布

二维均匀分布：设G是平面上的有界区域,面积为A,若二维随机变量(X,Y)具有概率密度

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

则称(X,Y)在G上服从均匀分布。若G_1是G内的面积为A_1的子区域，则有

$$P\{(X,Y)\in G_1\}=\iint\limits_{G_1}\frac{1}{A}dxdy=\frac{A_1}{A}$$

二维正态分布：设X,Y的联合密度为：

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

则称(X,Y)服从参数为$(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的二维正态分布，记为

$$(X,Y)\sim\mathbf{N}(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$$

其中各参数满足$-\infty<\mu<\infty,\sigma>0,|\rho|<1$，峰值$f(\mu_1,\mu_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$

二维正态分布的性质：二维正态r.v.的边际分布均为一维正态分布，若$(X,Y)\sim\mathbf{N}(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则$X\sim\mathbf{N}(\mu_1,\sigma_1^2),Y\sim\mathbf{N}(\mu_2,\sigma_2^2)$

边际密度均为正态分布的随机变量，其联合分布不一定是二维正态分布。

由随机变量的边际分布不能确定联合分布，反之可以。

连接函数(了解)：略

§3.4 独立随机变量

相互独立：设$(X,Y)\sim F(x,y)$，$X\sim F_X(x)$，$Y\sim F_Y(y)$，若$\forall x,y\in(-\infty,\infty)$有

$$P\{X\le x,Y\le y\}=P\{X\le x\}\cdot P\{Y\le y\}$$

即$F(x,y)=F_X(x)\cdot F_Y(y)$，则称随机变量$X,Y$相互独立

若(X,Y)相互独立，对于“合理”的集合$A,B\subset R$，$\{X\in A\},\{Y\in B\}$相互独立

离散r.v.相互独立：两个离散r.v.相互独立的充要条件是$p_{ij}=p_{i\cdot}\cdot p_{\cdot j}$

连续r.v.相互独立：$f(x,y)\overset{a.e.}{=}f_X(x)f_Y(y)$

若(X,Y)的密度函数能分解为$f(x,y)=g(x)h(y)$，$g(x)\ge 0$且$h(y)\ge 0$且支撑区域可分离，则X,Y相互独立

二维正态分布的性质：设$(X,Y)\sim\mathbf{N}(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则$X,Y$相互独立⇔$\rho=0$

n维随机变量的独立性：

1. $\forall x_1,x_2,\dots x_n\in R$，若有
   $F(x_1,x_2,\dots x_n)=F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$
   则称$X_1,X_2,\dots,X_n$相互独立
2. 设

$$\begin{gathered}(X_1,X_2,\cdots,X_m)\sim F_1(x_1,x_2,\cdots,x_m)\\(Y_1,Y_2,\cdots,Y_n)\sim F_2(y_1,y_2,\cdots,y_n)\\(X_1,X_2,\cdots,X_m,Y_1,Y_2,\cdots,Y_n)\sim F(x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_n)\end{gathered}$$

$\forall x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n \in R$，若有

$F(x_1, \ldots, x_m, y_1, \ldots, y_n) = F_1(x_1, \ldots, x_m) \cdot F_2(y_1, \ldots, y_n)$

则称$(X_1, \ldots, X_m)$，$(Y_1, \ldots, Y_n)$相互独立

3. 设$(X_1, \ldots, X_m)$，$(Y_1, \ldots, Y_n)$相互独立，则
   - X_i, Y_j相互独立
   - 设h, g分别是m元和n元的连续函数，则$h(X_1, \ldots, X_m), g(Y_1, \ldots, Y_n)$相互独立。

§3.5 条件分布

二维离散型r.v.的条件频率函数：设(X, Y)的频率函数为$P\{X = x_i, Y = y_j\} = p_{ij}$

考虑在$\{Y = y_j\}$已发生的条件下，$X = x_i$发生的条件概率$P\{X = x_i | Y = y_j\}$为

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

记为$P_{X|Y}(x_i|y_j)$，即为在$Y = y_j$的条件下，r.v. X的条件频率函数

条件频率函数的性质：

- 若X, Y独立，则$P_{X|Y}(x|y) = f_X(x)$
- $P\{X = x_i | Y = y_j\} \geq 0$
- $\sum\limits_{i=1}^{\infty} P\{X = x_i | Y = y_j\} = 1$，由上述两条可知条件频率函数也是频率函数
- 全概率公式：$f_X(x) = \sum\limits_y P_{X|Y}(x|y) f_Y(y)$

  若$f_Y(y) = 0$，则定义$P_{X|Y}(x|y) = 0$

二维r.v.的条件分布函数：$F_{X|Y}(x|y) = P\{X \leq x | Y = y\} = \frac{P\{X \leq x, Y=y\}}{P\{Y=y\}}$

- 对于连续型随机变量，若极限$\lim\limits_{\varepsilon \to 0^+} P\{X \leq x | y < Y \leq y + \varepsilon\}$存在，则称其为在条件$Y = y$下$X$的条件分布函数

二维连续型随机变量的条件概率密度：

$\lim\limits_{\varepsilon \to 0^+} P\{X \leq x | y < Y \leq y + \varepsilon\} = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$

条件密度：$f_{X|Y}(x|y) \triangleq \frac{f(x,y)}{f_Y(y)}$

条件分布：$F_{X|Y}(x|y) \triangleq \int_{-\infty}^{x} f_{X|Y}(u|y) du$

全概率公式 连续ver.：$f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X}(y|x) f_X(x) dx$

条件密度的性质：

- $f_{X|Y}(x|y) \geq 0$
- $\int_{-\infty}^{\infty} f_{X|Y}(x|y) dx = 1$