第二章：随机变量与分布函数

4. 随机向量的运算

4.1 加减乘除

4.2 线性变换

4.3 极值随机变量

4.4 一些特殊分布

1. 随机变量与分布函数

1.1 离散型随机变量与分布列

1.2 连续型随机变量与密度函数

1.3 一般型随机变量与分布函数

- 随机变量— Random Variable

In the late 1940s, William Feller (1907-1970) was working on his book *An introduction to probability theory and its applications, 1950, 1968*

While Joseph Doob (1910-2004) was working on his book *Stochastic Processes, 1953*

As Doob recalled: While working my book I had an argument with Feller. He asserted that everyone said "random variable" and I asserted that everyone said "chance variable". We obviously had to use the same name in our books, so we decided the issue by a stochastic procedure. That is, we tossed for it and he won.

Doob, 1910–2004

Feller, 1907-1970

- 随机变量
- 什么是随机变量?

随机变量是定义在概率空间上取实数值的可测函数

回忆：概率空间$(\Omega, \mathcal{A}, P)$是随机现象的数学描述

引入随机变量可以更好地使用现代数学工具分析随机现象

随机变量自然地出现在所讨论的随机现象中

(1) 掷骰子，记录出现的点数
(2) 测试灯泡的寿命
(3) 测量物体的长度

有时，很容易定义一个随机变量
(1) 抛硬币，出现正面，赢一元钱；出现反面，输一元钱
(2) 重复做试验，观察试验结果，记录成功的次数

根据随机变量取值的特点，可分为

(1) 离散型随机变量

(2) 连续型随机变量

(3) 一般型随机变量

不仅研究随机变量的取值，

而且研究取每个值的可能性大小—分布

分布（汉语词汇）:指在一定地区或区域内散布;

例句：翦伯赞(1898-1968)《内蒙访古》：“古城遗址最大多数分布在阴山南麓通向山北的峪口，也有分布在阴山北麓的。

1.1 离散型随机变量和分布列

取有限或可列个值的随机变量称为离散型随机变量

假定$(\Omega, \mathcal{A}, P)$是概率空间，

$$X : \Omega \longmapsto R$$

取值$x_1, x_2, \cdots, x_N$，$N < \infty$ 或$N = \infty$

取每个值的概率大小

$$P(X = x_i) = p_i, \quad i = 1, 2, \cdots, N$$

- 分布列：

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_N \\ p_1 & p_2 & \cdots & p_i & \cdots & p_N \end{pmatrix}$$

- 注：

(1)

$$p_i > 0, \quad \sum_{i=1}^{N} p_i = 1$$

(2) 对任意Borel 集B,

$$P(X \in B) = \sum_{i:x_i \in B} p_i$$

特别,

$$P(X \leq x) = \sum_{i:x_i \leq x} p_i$$

$$P(X > x) = \sum_{i:x_i > x} p_i$$

$$P(a < X \leq b) = \sum_{i:a < x_i \leq b} p_i$$

- 离散型随机变量的典型例子

1. 退化分布

$$X \sim \begin{pmatrix} c \\ 1 \end{pmatrix}$$

注：常数可看成是退化随机变量

2. 两点分布

$$X \sim \begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix}, \quad 0 < p < 1, p + q = 1$$

注：两点分布适用于描述“正面、反面”；“成功、失败”；“正常、维修”等随机现象

3. 二项分布

$$X \sim \begin{pmatrix} 0 & 1 & \cdots & k & \cdots & n \\ q^n & npq^{n-1} & \cdots & \binom{n}{k}p^k q^{n-k} & \cdots & p^n \end{pmatrix}$$

$0 < p < 1, p + q = 1$

简记$X \sim B(n,p)$

- 注：(1) 二项分布适用于n重Bernoulli试验。

记$P(A) = p$，X表示A发生的次数。

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \cdots, n$$

(2) 二项展开系数

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}$$

二项式系数表

- 在我国被称为贾宪三角或杨辉三角，一般认为是北宋数学家贾宪所首创。它记载于杨辉的《详解九章算法》(1261)之中。

- 在阿拉伯数学家卡西的著作《算术之钥》(1427)中也给出了一个二项式定理系数表，他所用的计算方法与贾宪的完全相同。

- 在欧洲，德国数学家阿皮安努斯在他1527年出版的算术书的封面上刻有此图。但一般却称之为帕斯卡三角形，因为帕斯卡在1654年也发现了这个结果。

(3) X最可能的值

$$
\begin{aligned}
\frac{p_k}{p_{k+1}} &= \frac{\binom{n}{k}p^k q^{n-k}}{\binom{n}{k+1}p^{k+1}q^{n-k+1}} \\
&= \frac{(k+1)q}{(n-k)p} \\
& \begin{cases} <1, & k+1<(n+1)p \\ >1, & k+1>(n+1)p \end{cases}
\end{aligned}
$$

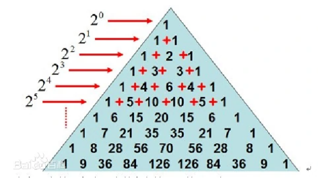
所以，

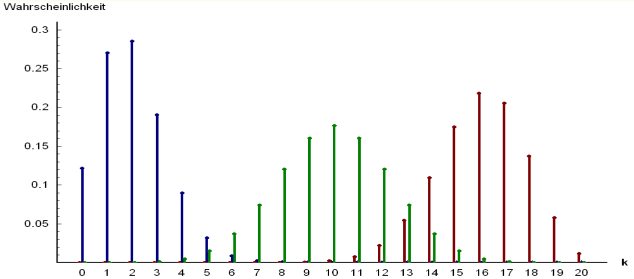
当$(n+1)p$ 是整数时，

$$p_{(n+1)p-1} = p_{(n+1)p}, \quad \text{达到最大}$$

当$(n+1)p$ 不是整数时，取整数部分$[(n+1)p]$

$$p_{[(n+1)p]}, \quad \text{达到最大}$$





Binomial distribution for $n = 20$

$p = 0.1$(blue), $p = 0.5$(green) and $p = 0.8$(red)

Binomial Calculator: Online Statistical Table

Use the Binomial Calculator to compute individual and cumulative binomial probabilities. For help in using the calculator, read the Frequently-Asked Questions or review the Sample Problems.

To learn more about the binomial distribution, go to Stat Trek's tutorial on the binomial distribution.

- **Enter a value in each of the first three text boxes (the unshaded boxes).**
- **Click the Calculate button.**
- **The Calculator will compute Binomial and Cumulative Probabilities.**

Probability of success on a single trial	
Number of trials	
Number of successes (x)	
Binomial probability: $P(X = x)$	
Cumulative probability: $P(X < x)$	
Cumulative probability: $P(X \leq x)$	
Cumulative probability: $P(X > x)$	
Cumulative probability: $P(X \geq x)$	

4. Poisson分布

$$X \sim \begin{pmatrix} 0 & 1 & \cdots & k & \cdots \\ e^{-\lambda} & \lambda e^{-\lambda} & \cdots & \frac{\lambda^k}{k!}e^{-\lambda} & \cdots \end{pmatrix}, \quad \lambda > 0$$

简记$X \sim \mathcal{P}(\lambda)$

X 取非负整数值，

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

注：(1)

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=1$$

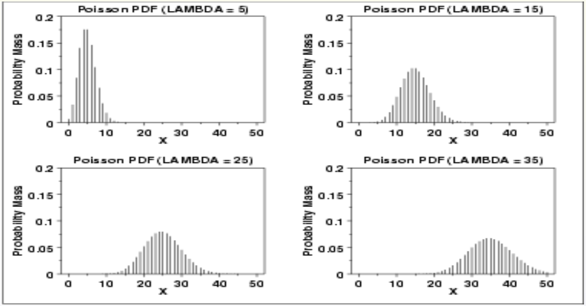
(2) Poisson分布是Poisson过程的基础，用于描述随机服务系统

(3) 当k足够大时，

$$\begin{aligned}P(X\geq k) &= \sum_{l=k}\frac{\lambda^l}{l!}e^{-\lambda}\\ &\asymp \frac{\lambda^k}{k!}e^{-\lambda}\end{aligned}$$

因此，X 只集中在小值上。

The following is the plot of the Poisson probability density function for four values of λ.



Poisson Distribution Calculator: Online Statistical Table

The Poisson Calculator makes it easy to compute individual and cumulative Poisson probabilities. For help in using the calculator, read the Frequently-Asked Questions or review the Sample Problems.

To learn more about the Poisson distribution, read Stat Trek's tutorial on the Poisson distribution.

- Enter a value in BOTH of the first two text boxes.
- Click the Calculate button.
- The Calculator will compute the Poisson and Cumulative Probabilities.

Poisson random variable (x)	
Average rate of success	
Poisson probability: P(X = x)	
Cumulative probability: P(X < x)	
Cumulative probability: P(X ≤ x)	
Cumulative probability: P(X > x)	
Cumulative probability: P(X ≥ x)	

- Poisson分布和二项分布之间有着下列关系
- Poisson 极限定理

假设$S_n \sim B(n, p_n)$。当$n \to \infty$，$np_n \to \lambda > 0$。那么对任意$k \geq 0$

$$P(S_n = k) \to \frac{\lambda^k}{k!} e^{-\lambda} = P(X = k), \quad n \to \infty$$

证明：固定k.

由于$S_n \sim B(n, p_n)$

$$\begin{aligned} P(S_n = k) &= \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} \\ &= \frac{1}{k!} \cdot n(n-1) \cdots (n-k+1) \cdot \frac{1}{n^k} \cdot (np_n)^k \\ &\quad \cdot (1 - \frac{\lambda}{n} + o(\frac{1}{n}))^{n-k} \\ &\to \frac{\lambda^k}{k!} e^{-\lambda}, \quad n \to \infty. \end{aligned}$$

例. Poisson approximation

Consider a video-poker plays who plays 1000 hands of five-card draw per day for 365 consecutive days. What is the probability that he is dealt a pat royal flush (皇家同花顺) exactly once; at least once?

We have $n = 365000$ independent Bernoulli trials, each with success probability

$$p = \frac{4}{\binom{52}{5}} = \frac{4}{2598960} = \frac{1}{649740}$$

Let $\lambda = np = \frac{365000}{649740}$. The exact binomial and approximation Poisson probabilities of exactly one success are

$$\binom{n}{1} p(1-p)^{n-1} \approx 0.320319294, \quad e^{-\lambda}\lambda \approx 0.320318939$$

For at least one success, they are

$$1 - \binom{n}{0} p(1-p)^{n-1} \approx 0.429797431, \quad 1 - e^{-\lambda}\lambda \approx 0.429797186$$

We see that the approximation is quite good!

- Siméon-Denis Poisson (21 June 1781 – 25 April 1840), was a French mathematician, geometer, and physicist.
- http://en.wikipedia.org/wiki/Sim

1781年6月21日生于法国皮蒂维耶，1840年4月25日卒于法国索镇。
泊松原读医科，1798年进巴黎综合工科学校改学数学，受到P.-S.拉普拉斯、J.-L.拉格朗日的赏识。1800年毕业后留校任教，1802年任副教授，1806年任教授。
1808年任法国经度局天文学家。1809年巴黎理学院成立，任该校数学教授。
1812年被选为法国科学院院士。

5. 几何分布

考虑随机试验E和事件A，$P(A)=p$，$0<p<1$。独立重复E，直至A发生，记录所做的试验次数，记为X.

- X 取正整数值，$1,2,\cdots$

$$P(X=k)=pq^{k-1}, \qquad \text{几何级数}$$

$$\sum_{k=1}^{\infty} pq^{k-1} = p\cdot\frac{1}{1-q} = 1.$$

- 分布列

$$X \sim \begin{pmatrix} 1 & 2 & \cdots & k & \cdots \\ p & pq & \cdots & pq^{k-1} & \cdots \end{pmatrix}, \quad 0<p<1$$

6. 超几何分布

考虑随机抽样。假设N件产品中含有M件次品，现随机抽样$n(\le N)$件，用X表示n件产品中次品数。

- $0 \le X \le \min\{M, n\}$

$$P(X=k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

- 证明

$$\sum_{k=0}^{\min\{M,n\}} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = 1 \quad \underline{\text{能直接证明吗？}}$$

- 几何级数序列$\{u_n, n \ge 1\}$: $\frac{u_{n+1}}{u_n} = q \ne 0$
- 超几何级数序列$\{v_n, n \ge 1\}$: $\frac{v_{n+1}}{v_n} = \frac{P(n)}{Q(n)}$，其中$P(n)$和$Q(n)$都是$n$的多项式。

1.2 连续型随机变量

- 连续型随机变量具有以下特点：

(1) 随机变量取值是一个或几个区间

(2) 存在一个函数$p(x)$

$$p(x) \geq 0, \quad \int_{-\infty}^{\infty} p(x)dx = 1$$

使得对任何Borel集B，

$$P(X \in B) = \int_B p(x)dx$$

简记$X \sim p(x)$

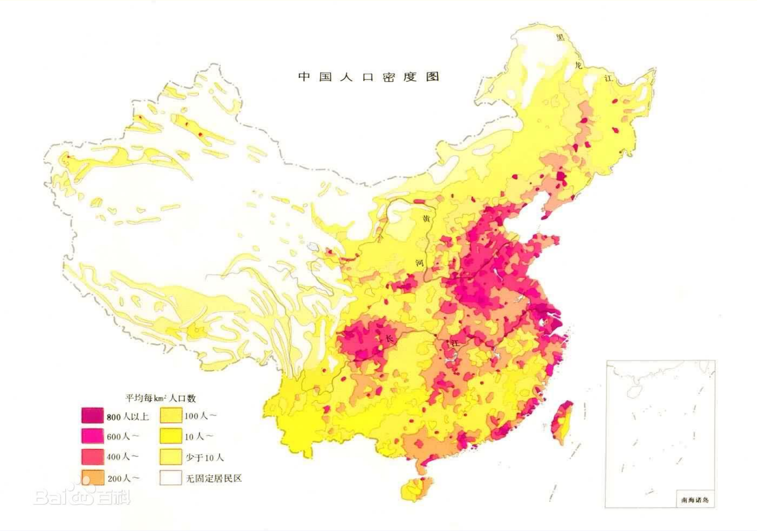
- 称$p(x)$为X的密度函数;
- 注:

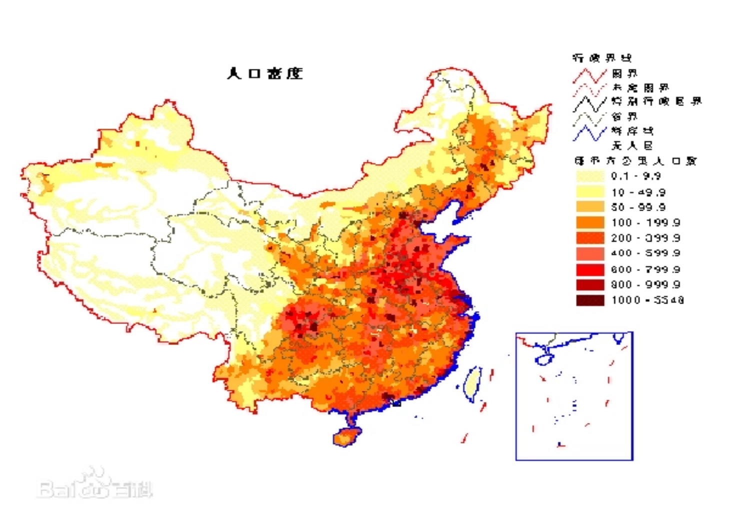
$$P(X = x) = 0, \quad x \in R$$

$$P(X \in (a, b]) = \int_a^b p(x)x, \quad a < b$$

- 物体质量密度: 密度是物质的特性之一，每种物质都有一定的密度，不同物质的密度一般是不同。因此我们可以利用密度来鉴别物质。

- 人口密度(density of population): 是单位土地面积上的人口数量。通常使用的计量单位有两种：人/平方公里;人/公顷。
它是衡量一个国家或地区人口分布状况的重要指标。


中国人口密度图
黑
龙
江
黄
河
长
江
平均每km² 人口数
800 人以上
600 人～
400 人～
200 人～
100 人～
10 人～
少于 10 人
无固定居民区
南海诸岛
Baidu百科


人口密度
行政界线
国界
未定国界
特别行政区界
省界
海岸线
无人区
每平方公里人口数
0.1 - 9.9
10 - 49.9
50 - 99.9
100 - 199.9
200 - 399.9
400 - 599.9
600 - 799.9
800 - 999.9
1000 - 5548

- 连续型随机变量的例子

1. 均匀分布

考虑随机试验：向(a,b)上随机投点，记落点的位置为X。那么

(1) X 落在(a,b)上每一点都是等可能的，

$$P(X=x)=0$$

(2) X 落在(a,b)上任何一个子区间的概率只与区间长度有关，与区间位置无关。一般地，

$$P(X\in A)=\frac{|A|}{b-a},\quad A\subseteq(a,b)\text{可测}$$

- X 是连续型随机变量，简记$X\sim U(a,b)$

$$X\sim p(x)=\begin{cases}\frac{1}{b-a}, & x\in(a,b)\\ 0, & \text{其他}\end{cases}$$

2. 指数分布

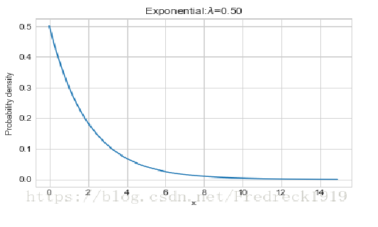
如果X取非负实数，且具有下列密度函数

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{其他} \end{cases} \qquad \lambda > 0$$

那么称X服从参数为λ的指数分布，简记$X \sim \exp(\lambda)$

- 注: (1)

$$p(x) \geq 0, \quad \int_{-\infty}^{\infty} p(x)dx = 1$$



(2) 指数分布通常用于描述人、零件的寿命；

(3) 指数分布与Poisson分布有着密切联系：Poisson 过程
(可参考《随机过程》苏中根编)

(4)

$$P(X > x) = e^{-\lambda x}, \quad x \geq 0$$

表明X取大值的可能性迅速衰减

(5)无记忆性

$$\begin{aligned} P(X > x + y | X > y) &= \frac{P(X > x + y)}{P(X > y)} \\ &= e^{-\lambda x} \\ &= P(X > x) \end{aligned}$$

6. 正态分布

假设随机变量X取所有实数值，并具有下列密度

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

那么称X是服从正态分布的随机变量，简记$X \sim N(\mu, \sigma^2)$

- 注：(1) $-\infty < \mu < \infty, \sigma > 0$

(2) 当$\mu = 0, \sigma^2 = 1$时，称$X \sim N(0,1)$ 是服从标准正态分布的随机变量

- 验证

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

经过变换，只要证明

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

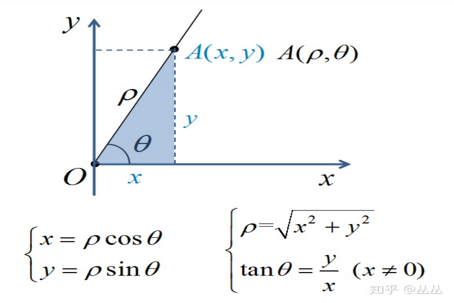
考虑

$$\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx\right)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dxdy$$

作极坐标变换：

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \end{cases}$$

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy &= \int_0^{\infty}\int_0^{2\pi}\frac{1}{2\pi}e^{-\frac{\rho^2}{2}}\rho d\rho d\theta \\ &= 1\end{aligned}$$



- 正态分布密度函数曲线具有良好的性质

(1) 对称性：关于$x = \mu$对称

(2) 光滑性：$p(x)$任意次可微

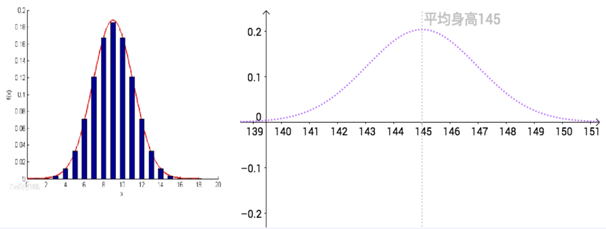
(3) 单调性：在$x = \mu$的左边，单调增加；在$x = \mu$的右边，单调减少

(4) $y = 0$是渐近线：当$x \to \pm\infty$时，$p(x) \to 0$

(5) 最大值: $P(x)$在$x = \mu$处取最大值$\frac{1}{\sqrt{2\pi}\sigma}$

(6) 倒钟形曲线

(7) σ变大，曲线变平坦；σ变小，曲线变陡峭



(8)

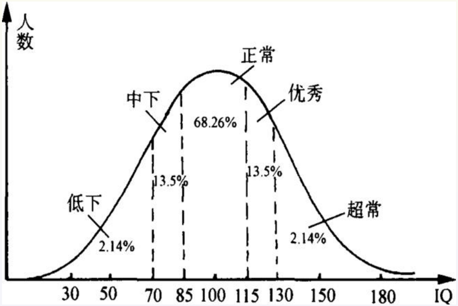
$$
\begin{aligned}
P(X > \mu + \sigma x) &= \int_{\mu+\sigma x}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du \\
&= \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \\
&\sim \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}}, \quad x \to \infty
\end{aligned}
$$

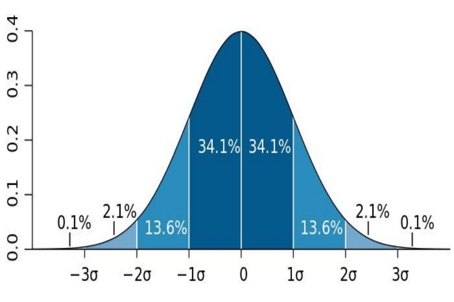
类似地，

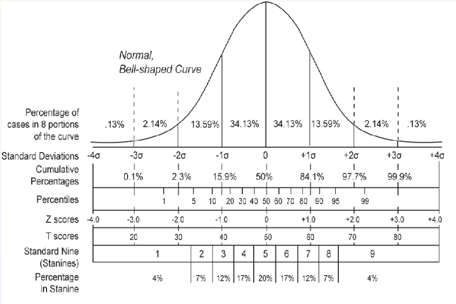
$$
P(X < \mu - \sigma x) \sim \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}}, \quad x \to \infty
$$

(9) X的取值集中在$[\mu-3\sigma,\mu+3\sigma]$里：$3\sigma$原则

$$P(|X-\mu|\leq 3\sigma)\asymp 0.9997$$




0.0
0.1
0.2
0.3
0.4
0.1%
2.1%
13.6%
34.1%
34.1%
13.6%
2.1%
0.1%
−3σ
−2σ
−1σ
0
1σ
2σ
3σ


Normal,
Bell-shaped Curve
Percentage of cases in 8 portions of the curve
.13%
2.14%
13.59%
34.13%
34.13%
13.59%
2.14%
.13%
Standard Deviations
-4σ -3σ -2σ -1σ 0 +1σ +2σ +3σ +4σ
Cumulative Percentages
0.1% 2.3% 15.9% 50% 84.1% 97.7% 99.9%
Percentiles
1 5 10 20 30 40 50 60 70 80 90 95 99
Z scores
-4.0 -3.0 -2.0 -1.0 0 +1.0 +2.0 +3.0 +4.0
T scores
20 30 40 50 60 70 80
Standard Nine (Stanines)
1 2 3 4 5 6 7 8 9
Percentage in Stanine
4% 7% 12% 17% 20% 17% 12% 7% 4%

1.3 一般随机变量

- 许多随机变量既不是离散型，也不是连续型随机变量。

例. X取非负实数，

$$P(X=0)=\frac{1}{2}$$

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

X 是混合型随机变量

- 随机变量

假定$(\Omega, \mathcal{A}, P)$是一概率空间，$(\mathbb{R}, \mathcal{B})$为Borel$\sigma$-域

$$\mathcal{B} = \sigma\Big\{ \sum_{i=1}^{m} (a_i, b_i], a_i < b_i < a_{i+1}, m \geq 1 \Big\}$$

函数

$$X : \Omega \mapsto \mathbb{R}$$

如果对任意Borel集$B \in \mathcal{B}$，

$$\{\omega : X(\omega) \in B\} \in \mathcal{A}$$

那么称X为随机变量

特别，

$$\{\omega : X(\omega) \leq x\}, \{\omega : X(\omega) > x\}, \{\omega : a < X(\omega) \leq b\} \in \mathcal{A}$$

- 分布函数

假定X是$(\Omega, \mathcal{A}, P)$上的随机变量。定义函数$F: \mathbb{R} \mapsto [0,1]$ 如下：

$$F(x) = P(\omega : X(\omega) \leq x), \quad x \in \mathbb{R}$$

称F为X的分布函数。

- 分布函数的性质

(1)

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \to \infty} F(x) = 1;$$

(2) $F(x)$ 单调不减:

$$F(x_1) \leq F(x_2), \quad x_1 < x_2$$

(3) $F(x)$ 左极限存在，右连续:

$$F(x-0) = \lim_{y \to x; y < x} F(y), \quad F(x+0) = \lim_{y \to x; y > x} F(y)$$

上述性质很容易从概率P的性质推导

具体地，

$$P(X<x)=F(x-0),\quad P(X=x)=F(x)-F(x-0)$$

$$P(X\in(a,b])=F(b)-F(a)$$

- 离散型随机变量的分布函数

假定

$$X\sim\begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_N \\ p_1 & p_2 & \cdots & p_i & \cdots & p_N \end{pmatrix}$$

那么

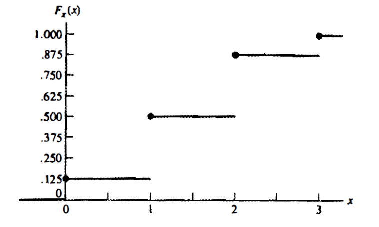
$$F(x)=\sum_{i:x_i\le x} p_i,\quad x\in\mathbb{R}$$

$F(x)$ 是一个阶梯型函数

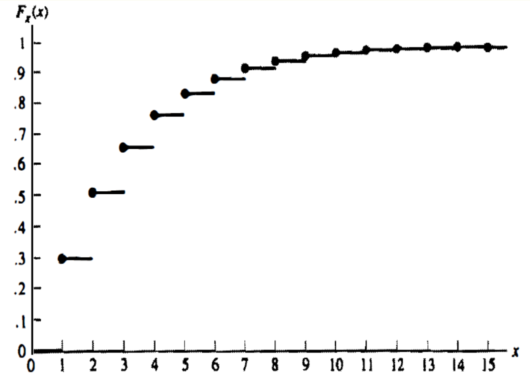
例1. 假定

$$X \sim \begin{pmatrix} -\frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{8} & \frac{1}{2} & \frac{3}{8} \end{pmatrix}$$

$$F(x) = \begin{cases} 0, & x < -\frac{1}{4} \\ \frac{1}{8}, & -\frac{1}{4} \leq x < 0 \\ \frac{5}{8}, & 0 \leq x < \frac{1}{2} \\ 1, & \frac{1}{2} \leq x \end{cases}$$



例2. 考虑几何随机变量X, 参数为$p = 0.3$.



- 连续型随机变量的分布函数

假定X是连续型随机变量，$X \sim p(x)$。那么

$$
\begin{aligned}
F(x) &= P(X \leq x) \\
&= \int_{-\infty}^{x} p(u)du
\end{aligned}
$$

$F(x)$是连续函数，并且有导数

$$F'(x) = p(x)$$

例1. 均匀分布的分布函数

假定$X \sim U(a,b)$，那么

$$
\begin{aligned}
F(x) &= \int_{-\infty}^{x} p(u)du \\
&= \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x < b \\ 1, & x \ge b \end{cases}
\end{aligned}
$$

$F(x)$在(a,b)内线性增长

例2. 指数分布的分布函数

假定$X \sim \exp(\lambda)$，那么

$$
\begin{aligned}
F(x) &= \int_{-\infty}^{x} p(u)du \\
&= \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}
\end{aligned}
$$

例3. 正态分布的分布函数

假定$X \sim N(0,1)$，那么

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

- 没有显性表达式。
- 当随机变量服从标准正态分布时，通常用$\Phi(x)$ 表示其分布函数。在机器学习中，称$\Phi(x)$为Probit 函数
- 有时，采用

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-\frac{u^2}{2}} du, \qquad \text{误差函数}$$

$$\Phi(x) = \frac{1}{2}\{1 + \frac{1}{\sqrt{2}} erf(x)\}$$

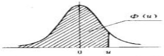
- $\Phi(x)$的值可以通过查表得到. 如

$$\Phi(1) = 0.841345, \quad \Phi(2) = 0.977250, \quad \Phi(3) = 0.998650$$

GB 4086.1—83

2 正态分布函数表

$$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$



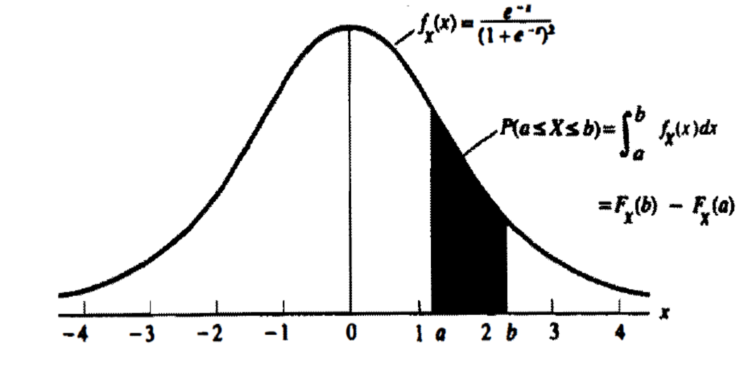
u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500000	0.503989	0.507978	0.511966	0.515953	0.519939	0.523922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555670	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621720	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802337	0.805105	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823814	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.879000	0.881000	0.882977
1.2	0.884930	0.886861	0.888768	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903200	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935745	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.959070	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.964070	0.964852	0.965620	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705
2.0	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999
3.1	0.999032	0.999065	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289
3.2	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499
3.3	0.999517	0.999534	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999651
3.4	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758
3.5	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999822	0.999828	0.999835
3.6	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888
3.7	0.999892	0.999896	0.999900	0.999904	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925
3.8	0.999928	0.999931	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950
3.9	0.999952	0.999954	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967
4.0	0.999968	0.999970	0.999971	0.999972	0.999973	0.999974	0.999975	0.999976	0.999977	0.999978
4.1	0.999979	0.999980	0.999981	0.999982	0.999983	0.999983	0.999984	0.999985	0.999985	0.999986
4.2	0.999987	0.999987	0.999988	0.999988	0.999989	0.999989	0.999990	0.999990	0.999991	0.999991
4.3	0.999991	0.999992	0.999992	0.999993	0.999993	0.999993	0.999993	0.999994	0.999994	0.999994
4.4	0.999995	0.999995	0.999995	0.999995	0.999996	0.999996	0.999996	0.999996	0.999996	0.999996
4.5	0.999997	0.999997	0.999997	0.999997	0.999997	0.999997	0.999997	0.999998	0.999998	0.999998
4.6	0.999998	0.999998	0.999998	0.999998	0.999998	0.999998	0.999998	0.999998	0.999999	0.999999
4.7	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999
4.8	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999
4.9	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

本表对于u给出正态分布函数$\Phi(u)$的数值。

例：对于$u=1.33$，$\Phi(u)=0.908241$。

例4. Logistic 密度函数和分布函数:

$$X \sim p(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \quad F_X(x) = \frac{1}{1+e^{-x}}, \quad x \in \mathbb{R}$$



- 回忆分布函数的性质

(1)

$$F(-\infty)=\lim_{x\to-\infty}F(x)=0,\quad F(+\infty)=\lim_{x\to\infty}F(x)=1;$$

(2) $F(x)$ 单调不减:

$$F(x_1)\le F(x_2),\quad x_1<x_2$$

(3) $F(x)$ 左极限存在，右连续:

$$F(x-0)=\lim_{y\to x;y<x}F(y),\quad F(x+0)=\lim_{y\to x;y>x}F(y)$$

- 通常具有上述(1), (2), (3)的实函数F为分布函数
- 给定这样的函数F，一定能找到一个概率空间$(\Omega,\mathcal{A},P)$和一个随机变量$X:\Omega\mapsto\mathbb{R}$，使得

$$F(x)=P(X\le x)$$

2. 随机变量的函数

前一节，介绍了一些简单随机变量

但实际问题中通常需要考虑复杂随机变量，它们大多是上述简单随机变量的函数

如，

(1) 赌博往往更关心输赢钱数和赌资，投掷骰子或硬币只是工具

$$H \Rightarrow +1\text{元}, \quad T \Rightarrow -1\text{元}$$

(2) 如何收取保费依赖于寿命长短，并不是简单的比例关系:

保险公司		百年人寿
产品名称		**康多保**
投保条件	投保年龄	28天-60周岁
	保障期限	终身
	最长缴费期限	30年
	等待期	90天
重症保障	种类	100种
	保额	100%保额
	次数	5次（分5组）
中症保障	种类	20种
	保额	60%保额
	次数	2次（不分组）
轻症保障	种类	35种
	保额	依次赔付35%、40%、45%保额
	次数	3次（不分组）
其他保障	被保人豁免	轻症、中症、重疾
	身故/全残	18周岁前，已交保费的3倍； 18周岁后，100%保额
保费测算		50万保额，保终身，30年交费
30岁男		9360元
30岁女		8640元

一般地，人们需要考虑

$$Y = f(X), \quad Y(\omega) = f(X(\omega))$$

其中X 是一个已知的随机变量，f是一个函数

- 问题：

(1) Y确实是一个随机变量吗？

(2) 如何计算Y的分布？

- (1) 什么条件下Y 是一个随机变量吗？

当f是可测函数时，Y是一个随机变量

$$B\text{可测} \Longrightarrow f^{-1}(B) = \{x : f(x) \in B\}\text{可测}$$

$$\{Y \in B\} = \{f(X) \in B\}$$

$$\{f(X) \in B\} = \{X \in f^{-1}(B)\} \in \mathcal{A}$$

- 常见的函数都是可测的

如，

(1) 连续函数

(2) 分段连续函数

(3) 单调函数

(4) 分段单调函数

(2) 如何计算Y的分布？

- 当X是离散型随机变量时，Y仍是离散型随机变量。

其取值和分布可以通过直接计算得到

例1. 假定

$$X \sim \begin{pmatrix} -1 & 0 & 1 & 2 \\ \frac{1}{8} & \frac{1}{2} & \frac{1}{8} & \frac{1}{4} \end{pmatrix}$$

求$Y = X^2$的分布列

$$Y \sim \begin{pmatrix} 1 & 0 & 4 \\ \frac{1}{8} + \frac{1}{8} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

一般地，假定

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_N \\ p_1 & p_2 & \cdots & p_i & \cdots & p_N \end{pmatrix}$$

$$Y = f(X)$$

(1)改写

$$\{f(x_1), f(x_2), \cdots, f(x_N)\} = \{y_1, y_2, \cdots, y_k\}$$

$y_1, y_2, \cdots, y_k$ 互不相同

(2)

$$P(Y = y_l) = \sum_{i: f(x_i) = y_l} p_i$$

- 当X是连续型随机变量时，Y不一定是连续型随机变量。

依赖于函数f的性质，没有统一的计算公式

例. 假设$X \sim N(0,1)$，

$$f(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases}$$

求$Y = f(X)$的分布？

$$Y \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

例2. 假设$X \sim \exp(\lambda)$,

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \log x, & x > 0 \end{cases}$$

求$Y = f(X)$的分布?

对任意$-\infty < y < \infty$,

$$\begin{aligned} P(Y \leq y) &= P(\log X \leq y) \\ &= P(X \leq e^y) \\ &= 1 - e^{-\lambda e^y} \end{aligned}$$

Y是连续型随机变量

例3. 假设$X \sim N(0,1)$，求$Y=|X|$的分布？

Y只取非负实数值。对任意$y \geq 0$，

$$
\begin{aligned}
P(Y \leq y) &= P(|X| \leq y) \\
&= P(-y \leq X \leq y) \\
&= P(X \leq y) - P(X \leq -y) \\
&= \Phi(y) - \Phi(-y)
\end{aligned}
$$

Y是连续型随机变量，具有密度函数

$$
Y \sim \Phi'(y) + \Phi'(-y) = \frac{\sqrt{2}}{\sqrt{\pi}} e^{-\frac{y^2}{2}}, \quad y \geq 0
$$

例4. 假设$X \sim N(\mu, \sigma^2)$，求$Y = \frac{X-\mu}{\sigma}$的分布?

对任意$-\infty < y < \infty$，

$$
\begin{aligned}
P(Y \leq y) &= P(\frac{X-\mu}{\sigma} \leq y) \\
&= P(X \leq \mu + \sigma y) \\
&= \int_{-\infty}^{\mu+\sigma y} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du \\
&= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du
\end{aligned}
$$

- $Y \sim N(0, 1)$

一般地，

- 假设$X \sim p_X(x)$，$f = f(x)$具有反函数f^{-1}，并且f^{-1}可导。那么$Y = f(X)$仍是连续型随机变量，具有密度函数

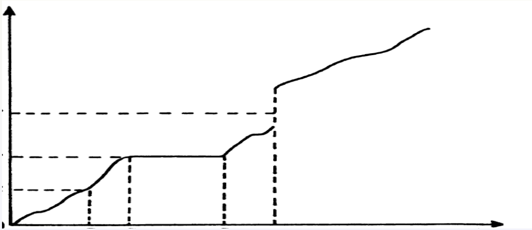
$$Y \sim p_Y(y) = p_X(f^{-1}(y))|(f^{-1})'(y)|$$

其中y位于f的值域中

例1. 假设$X \sim U(0,1)$，$F(x)$是给定的分布函数，求$Y = F^{-1}(X)$分布？

注意，$F(x)$具有良好的性质：$0 \leq F(x) \leq 1$，单调不减，右连续. 定义

$$F^{-1}(y) = \sup\{x : F(x) < y\}$$



根据上确界的定义，可以验证

(1) $F^{-1}(y)(0 < y < 1)$ 是y的单调不减函数；

(2) $F(F^{-1}(y)) \leq y$.

如果$F(x)$ 在$x = F^{-1}(y)$处连续，则$F(F^{-1}(y)) = y$

(3) $F^{-1}(y) \leq x$ 当且仅当$y \leq F(x)$.

这样，我们有

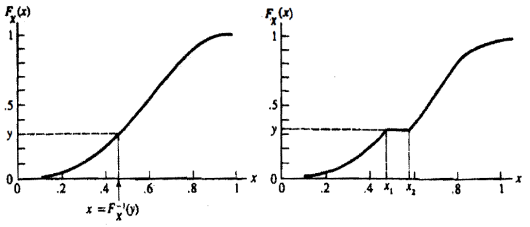
$$\begin{aligned} P(Y \leq y) &= P(F^{-1}(X) \leq y) \\ &= P(X \leq F(y)) = F(y) \end{aligned}$$

- $Y \sim F(x)$

例2. 假设X具有连续的分布函数$F(x)$，求$Y=F(X)$的分布？

$$
\begin{aligned}
P(Y\le y) &= P(F(X)\le y)\\
&= P(X\le F^{-1}(y))\\
&= F(F^{-1}(y)) \overset{\text{连续性}}{=} y,\quad 0\le y\le 1
\end{aligned}
$$

- $Y\sim U(0,1)$



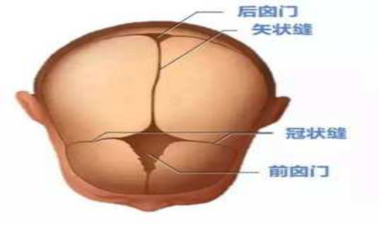
3. 随机向量和联合分布函数

- 人们经常需要研究和使用随机向量

(1) 抽查高考考生成绩

名	语文	数学	英语	理综	总分
谦	128	135	140	276	679
靖	126	138	133	281	678
夫	129	128	136	278	671
扬	128	142	136	263	669
元	128	136	138	266	668
童	123	144	139	258	664
友	118	142	135	268	663
诗	131	138	131	263	663
彬	115	136	131	277	659
涛	129	128	145	257	659

(2) 婴儿体检: 量身高、测体重、看囟门(xin men) 、测胸围、测头围



(3) 飞行器在空中的位置

一般地，给定概率空间$(\Omega, \mathcal{A}, P)$，

$$\mathbf{X}: \Omega \mapsto R^m, \qquad \mathbf{X}(\omega) = (X_1(\omega), X_2(\omega)\cdots, X_m(\omega))$$

其中，$X_1, X_2, \cdots, X_m$都是随机变量, 称$\mathbf{X}$为m-维随机向量

- 对于随机向量$\mathbf{X}$，

不仅需要研究每个分量的分布，

而且需要研究各个分量之间的关系，即联合分布

为便于理解概念，我们着重介绍2-随机向量

高维情形可以类似讨论

3.1 离散型随机向量和分布表

正如以前一样，我们首先介绍离散型随机向量。

- 给定概率空间$(\Omega, \mathcal{A}, P)$，(X, Y)是2-随机向量。

假设X取值为$x_1, x_2, \cdots$；Y取值为$y_1, y_2, \cdots$。那么称(X, Y)为离散型随机向量

记

$$p_{ij} = P(X = x_i, Y = y_j), \quad i, j = 1, 2, \cdots$$

$$p_{ij} \geq 0, \quad \sum_{i,j} p_{ij} = 1$$

称

$$((x_i, y_j), p_{ij})_{i,j=1}^{\infty}$$

为随机向量(X, Y)的联合分布。

对任意Borel集A, B，

$$P(X \in A, Y \in B) = \sum_{i:x_i \in A, j:y_j \in B} p_{ij}$$

特别，

$$P(X \leq x, Y \leq y) = \sum_{i:x_i \leq x, j:y_j \leq y} p_{ij}$$

- 边际分布

X,Y的分布可以由p_{ij}计算得到

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_{1\cdot} & p_{2\cdot} & \cdots & p_{i\cdot} & \cdots \end{pmatrix}$$

其中

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

$$p_{i\cdot} > 0, \quad \sum_{i=1}^{\infty} p_{i\cdot} = 1$$

类似地，

$$Y \sim \begin{pmatrix} y_1 & y_2 & \cdots & y_j & \cdots \\ p_{\cdot 1} & p_{\cdot 2} & \cdots & p_{\cdot j} & \cdots \end{pmatrix}$$

其中

$$p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij}, \quad p_{\cdot j} > 0, \quad \sum_{j=1}^{\infty} p_{\cdot j} = 1$$

$\xi \backslash \eta$	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{11}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

- 边际分布由联合分布所惟一确定

但边际分布不能惟一决定联合分布

例1. 一罐子里装有2个白球，3个黑球。现随机依次抽取两球，每次一个。分别用X, Y表示第一、二次取得白球的个数。求(X, Y)的联合分布？

显然，X, Y都只取$0, 1$两个值。向量(X, Y)取4对值：

$$(0,0), \quad (1,0), \quad (0,1), \quad (1,1)$$

下面求取每对值的联合分布。 按(1) 放回, (2) 无放回进行讨论。

- (1) 放回

$$P(X=0,Y=0)=P(X=0)\cdot P(Y=0|X=0)=\frac{3}{5}\cdot\frac{3}{5}$$

$$P(X=0,Y=1)=P(X=0)\cdot P(Y=1|X=0)=\frac{3}{5}\cdot\frac{2}{5}$$

$$P(X=1,Y=0)=P(X=1)\cdot P(Y=0|X=1)=\frac{2}{5}\cdot\frac{3}{5}$$

$$P(X=1,Y=1)=P(X=1)\cdot P(Y=1|X=1)=\frac{2}{5}\cdot\frac{2}{5}$$

- (2) 无放回

$$P(X=0,Y=0)=P(X=0)\cdot P(Y=0|X=0)=\frac{3}{5}\cdot\frac{2}{4}$$

$$P(X=0,Y=1)=P(X=0)\cdot P(Y=1|X=0)=\frac{3}{5}\cdot\frac{2}{4}$$

$$P(X=1,Y=0)=P(X=1)\cdot P(Y=0|X=1)=\frac{2}{5}\cdot\frac{3}{4}$$

$$P(X=1,Y=1)=P(X=1)\cdot P(Y=1|X=1)=\frac{2}{5}\cdot\frac{1}{4}$$

下面求X, Y的边际分布

- (1) 放回

$$X \sim \begin{pmatrix} 0 & 1 \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix}$$

- (2)无放回

$$X \sim \begin{pmatrix} 0 & 1 \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix}$$

- 条件分布

第一章学习了事件的条件概率。对于固定的x_i, y_j，我们得到

$$\begin{aligned} P(Y = y_j | X = x_i) &= \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} \\ &= \frac{p_{ij}}{p_{i\cdot}} \end{aligned}$$

如果x_i, y_j变化，那么我们可以得到<u>条件分布</u>:

给定$X = x_i$的条件下，Y可取值$y_1, y_2, \cdots, y_j, \cdots$

概率分别为

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

- 条件分布列

$$Y | X = x_i \sim \begin{pmatrix} y_1 & y_2 & \cdots & y_j & \cdots \\ \frac{p_{i1}}{p_{i\cdot}} & \frac{p_{i2}}{p_{i\cdot}} & \cdots & \frac{p_{ij}}{p_{i\cdot}} & \cdots \end{pmatrix}$$

类似地，给定$Y=y_j$的条件下，X可取值$x_1, x_2, \cdots, x_i, \cdots$概率分别为

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

$$X|Y=y_j \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ \frac{p_{1j}}{p_{\cdot j}} & \frac{p_{2j}}{p_{\cdot j}} & \cdots & \frac{p_{ij}}{p_{\cdot j}} & \cdots \end{pmatrix}$$

- 独立随机变量(离散型)

假设(X, Y)是离散型随机向量(分布表如上)。如果

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j), \qquad \forall i, j$$

即

$$p_{ij} = p_{i\cdot} \cdot p_{\cdot j}, \qquad \forall i, j$$

那么称X和Y相互独立

注：(1) 随机变量的独立型不同于事件的独立。它要求对所有的i, j，事件$X = x_i$和事件$Y = y_j$相互独立
(2) 如果存在一对i, j使得

$$P(X = x_i, Y = y_j) \neq P(X = x_i)P(Y = y_j)$$

那么X和Y不相互独立

例1. 随机向量(X,Y)分布如下：

$$P(X=1,Y=1)=\frac{1}{6},\quad P(X=2,Y=1)=\frac{1}{3}$$
$$P(X=1,Y=2)=\frac{1}{9},\quad P(X=2,Y=2)=a$$
$$P(X=1,Y=3)=\frac{1}{18},\quad P(X=2,Y=3)=b$$

(1) 各概率之和为1

$$\frac{1}{6}+\frac{1}{3}+\frac{1}{9}+a+\frac{1}{18}+b=1$$

(2) 边际分布分别为

$$X\sim\begin{pmatrix} 1 & 2 \\ \frac{1}{6}+\frac{1}{9}+\frac{1}{18} & \frac{1}{3}+a+b \end{pmatrix}$$

$$Y\sim\begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{6}+\frac{1}{3} & \frac{1}{9}+a & \frac{1}{18}+b \end{pmatrix}$$

(3) 独立性条件

$$P(X=1,Y=2)=P(X=1)P(Y=2)$$

$$P(X=1,Y=3)=P(X=1)P(Y=3)$$

即

$$\frac{1}{9}=\frac{1}{3}\cdot(\frac{1}{9}+a),\qquad a=\frac{2}{9}$$

$$\frac{1}{18}=\frac{1}{3}\cdot(\frac{1}{18}+b),\qquad b=\frac{1}{9}$$

将$a=\frac{2}{9},b=\frac{1}{9}$代入联合分布中，验证其他独立性条件

3.2 连续型随机向量

给定概率空间$(\Omega, \mathcal{A}, P)$，(X, Y)是其上的随机向量。如果存在$p(x, y)$使得

$$p(x, y) \geq 0, \quad \int_{-\infty}^{\infty} p(x, y) dx dy = 1$$

并且对任意Borel 集$A, B \subset R$，

$$P(X \in A, Y \in B) = \int_A \int_B p(u, v) du dv$$

那么称(X, Y)是连续型随机向量，具有密度函数$p(x, y)$

特别，对任意$x, y \in R$

$$P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p(u, v) du dv$$

- 边际分布

显然, 如果(X,Y)是连续型随机向量，那么X,Y 都是连续型随机变量。

并且，

$$X \sim p_X(x), \quad Y \sim p_Y(y)$$

其中

$$p_X(x) = \int_{-\infty}^{\infty} p(x,y)dy, \quad p_Y(y) = \int_{-\infty}^{\infty} p(x,y)dx$$

同样，联合分布惟一决定边际分布；边际分布不能决定联合分布

例1. 矩形区域上的均匀分布 假设$(a,b)\times(c,d)$是一个矩形区域，如果随机向量(X,Y)具有密度函数

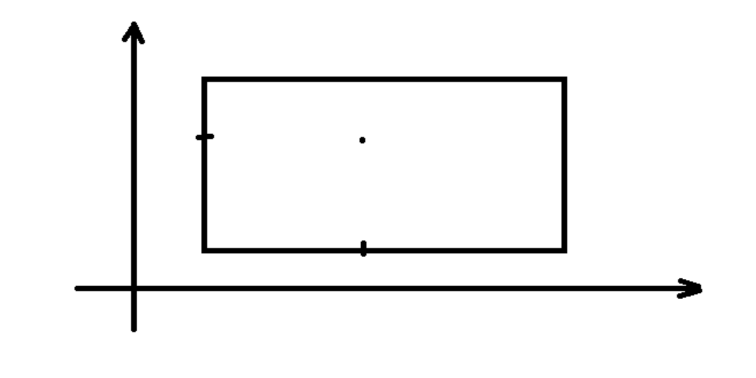
$$p(x,y)=\begin{cases}\frac{1}{(b-a)(d-c)} & (x,y)\in(a,b)\times(c,d)\\ 0 & \text{其它}\end{cases}$$

那么(X,Y)是$(a,b)\times(c,d)$上的均匀分布

注：

(1) (X,Y)落在$(a,b)\times(c,d)$上每一点都是等可能的；

(2) 落在子区域G上的可能性只与G的面积大小成比例，与位置无关。



如果(X,Y)是$(a,b)\times(c,d)$上的均匀分布，那么

$$
\begin{aligned}
P(X\le x) &= P(X\le x, Y<\infty)\\
&= \int_{-\infty}^{x}\int_{-\infty}^{\infty} p(u,v)\,dudv\\
&= \begin{cases} 0 & x<a \\ \frac{x-a}{(b-a)} & a\le x<b \\ 1 & x\ge b \end{cases}
\end{aligned}
$$

所以，X是(a,b)上的均匀分布

类似地，Y是(c,d)上的均匀分布

例2. 单位圆上的均匀分布

假设$B((0,0),1)$是中心在$(0,0)$处的单位圆。如果随机向量(X,Y)具有密度函数

$$p(x,y)=\begin{cases}\frac{1}{\pi} & (x,y)\in B((0,0),1)\\ 0 & \text{其它}\end{cases}$$

那么称(X,Y)为单位圆上的均匀分布

注：

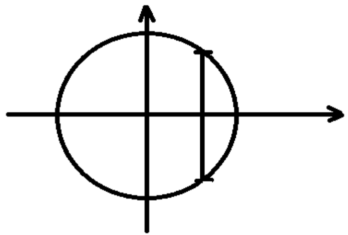
(1) (X,Y)落在$B((0,0),1)$上每一点都是等可能的；

(2) 落在子区域G上的可能性只与G的面积大小成比例，与位置无关。

- 边际分布

X的分布

$$
\begin{aligned}
X \sim p_X(x) &= \int_{-\infty}^{\infty} p_{x,y}\,dy \\
&= \begin{cases} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dy & -1 < x < 1 \\ 0 & \text{其它} \end{cases} \\
&= \begin{cases} \frac{2\sqrt{1-x^2}}{\pi} & -1 < x < 1 \\ 0 & \text{其它} \end{cases}
\end{aligned}
$$



类似地，Y的分布

$$
\begin{aligned}
Y \sim p_Y(y) &= \int_{-\infty}^{\infty} p_{x,y} dx \\
&= \begin{cases} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} dx & -1 < y < 1 \\ 0 & \text{其它} \end{cases} \\
&= \begin{cases} \frac{2\sqrt{1-y^2}}{\pi} & -1 < y < 1 \\ 0 & \text{其它} \end{cases}
\end{aligned}
$$

注：例1, 2尽管联合都是均匀分布，但边际分布完全不同

例3. 联合正态分布

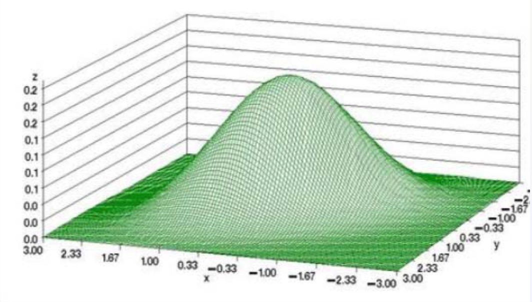
如果随机向量(X,Y)具有密度函数: $\forall x,y\in R$

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

那么称(X,Y)服从二元联合正态分布，$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$为参数.

简记

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



•验证$p(x,y)$确实是密度函数

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

经过变换: $u=\frac{x-\mu_1}{\sigma_1}$, $v=\frac{y-\mu_2}{\sigma_2}$，等价于

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[x^2-2\rho xy+y^2\right]}dxdy=1$$

左边进行配方，得

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}(x-\rho y)^2}e^{-y^2/2}dxdy$$

$$=\int_{-\infty}^{\infty}\left(\underline{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi(1-\rho^2)}}e^{-\frac{1}{2(1-\rho^2)}(x-\rho y)^2}dx}\right)\cdot\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy$$

$$=\int_{-\infty}^{\infty}1\cdot\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy=1$$

- 边际分布

假设$(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2;\rho)$，求$X,Y$的边际分布?

$$X\sim p_X(x)$$

$$\begin{aligned}
p_X(x) &= \int_{-\infty}^{\infty} p(x,y)dy \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\
&\quad \times e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{2\sigma_2^2}\right]}dy
\end{aligned}$$

进行配方得,

$$\begin{aligned} p_X(x) &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} \\ &\quad \times e^{-\frac{1}{2(1-\rho^2)\sigma_2^2}\left(y-\mu_2-\frac{\sigma_2\rho(x-\mu_1)}{\sigma_1}\right)^2} dy \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \end{aligned}$$

因此,

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

类似地,

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

- 二元正态分布的边际分布仍是正态分布

- 条件分布

假设(X,Y)是连续型随机向量，具有联合密度函数$p(x,y)$。下面讨论条件分布

给定$X=x$，求Y的分布？

即$P(Y\le y|X=x)$

注意$P(X=x)=0$，我们采用

$$P(Y\le y|X=x)=\lim_{\varepsilon\to 0}P(Y\le y|x-\varepsilon<X\le x+\varepsilon)$$

$$\begin{aligned}
P(Y \le y|X=x) &= \lim_{\varepsilon\to 0}\frac{P(Y\le y, x-\varepsilon<X\le x+\varepsilon)}{P(x-\varepsilon<X\le x+\varepsilon)} \\
&= \lim_{\varepsilon\to 0}\frac{P(x-\varepsilon<X\le x+\varepsilon, Y\le y)}{P(x-\varepsilon<X\le x+\varepsilon)} \\
&= \lim_{\varepsilon\to 0}\frac{\int_{x-\varepsilon}^{x+\varepsilon}\int_{-\infty}^{y}p(u,v)dudv/(2\varepsilon)}{\int_{x-\varepsilon}^{x+\varepsilon}p_X(u)du/(2\varepsilon)} \\
&= \frac{\int_{-\infty}^{y}p(x,v)dv}{p_X(x)}
\end{aligned}$$

给定$X=x$下，Y具有密度函数

$$p_{Y|X}(y|x)=\frac{p(x,y)}{p_X(x)}$$

类似地，

给定$Y=y$下，X具有密度函数

$$p_{X|Y}(x|y)=\frac{p(x,y)}{p_Y(y)}$$

考虑$(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2;\rho)$。求$p_{X|Y}(x|y),p_{Y|X}(y|x)$?

$$\begin{aligned}
p_{X|Y}(x|y) &= \frac{p(x,y)}{p_Y(y)} \\
&= \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}}{\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}}
\end{aligned}$$

进行整理得,

$$p_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} e^{-\frac{\left(x-\mu_1-\frac{\rho\sigma_1}{\sigma_2}(y-\mu_2)\right)^2}{2\sigma_1^2(1-\rho^2)}}$$

即固定y,

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

类似地，固定x,

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

- 独立随机变量(连续型)

假设$(X,Y)\sim p(x,y)$，$X\sim p_X(x), Y\sim p_Y(y)$。
如果

$$p(x,y)=p_X(x)p_Y(y),\quad \forall x,y$$

那么称X,Y相互独立

回忆，假设(X,Y)是离散型随机向量，具有分布p_{ij}.
如果

$$p_{ij}=p_{i\cdot}p_{\cdot j},\quad \forall i,j$$

那么称X,Y相互独立

例1. 如果(X,Y)是矩形区域$(a,b)\times(c,d)$上的均匀分布，那么X,Y相互独立

例2. 如果(X,Y)是单位圆$B((0,0),1)$上的均匀分布，那么X,Y不相互独立

例3. 如果$(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2;\rho)$，那么
(1) 当$\rho=0$时，X,Y相互独立
(2) 当$\rho\neq 0$时，X,Y不相互独立

- 一般随机向量和联合分布函数
- 存在随机向量(X,Y)，既不是离散型，也不是连续型的。

例. 考虑随机向量(X,Y)，其中$X\sim U(0,1)$, $Y\sim B(n,X)$。

- 对于一般的随机向量(X,Y)，我们主要利用其分布函数

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

来进行研究

注：

(1) 如果(X,Y)是离散型，那么

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

(2) 如果(X,Y)是连续型，那么

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

- $F(x,y)$具有下列基本性质：

(1)

$$F(-\infty, y) = F(x, -\infty) = 0, \quad F(+\infty, +\infty) = 1$$

(2) $F(x,y)$关于x和y单调不减

(3) $F(x,y)$关于x和y左极限存在，右连续

(4)

$$P(a < X \leq b, c < Y \leq d) \quad = \quad F(a,c) + F(b,d) - F(a,d) - F(b,c)$$

上述性质可以从概率的基本性质推出

- 边际分布

假设(X,Y)具有分布函数$F(x,y)$，那么X,Y的分布函数分别为

$$
\begin{aligned}
F_X(x) &= P(X \leq x) \\
&= P(X \leq x, Y < +\infty) \\
&= F(x, +\infty)
\end{aligned}
$$

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(X < +\infty, Y \leq y,) \\
&= F(+\infty, y)
\end{aligned}
$$

- 条件分布

假设(X,Y)具有分布函数$F(x,y)$，那么给定$X=x$, Y的条件分布函数为

$$
\begin{aligned}
P(Y\le y|X=x) &= \lim_{\varepsilon\to 0} P(Y\le y|x-\varepsilon<X\le x+\varepsilon) \\
&= \lim_{\varepsilon\to 0}\frac{F(x+\varepsilon,y)-F(x-\varepsilon,y)}{F_X(x+\varepsilon)-F_X(x-\varepsilon)}
\end{aligned}
$$

类似地，

$$
\begin{aligned}
P(X\le x|Y=y) &= \lim_{\varepsilon\to 0} P(X\le x|y-\varepsilon<Y\le y+\varepsilon) \\
&= \lim_{\varepsilon\to 0}\frac{F(x,y+\varepsilon)-F(x,y-\varepsilon)}{F_Y(y+\varepsilon)-F_Y(y-\varepsilon)}
\end{aligned}
$$

- 独立随机向量

假设(X,Y)具有分布函数$F(x,y)$，边际分布为$F_X(x)$,$F_Y(y)$. 如果

$$F(x,y)=F_X(x)F_Y(y),\quad \forall x,y\in\mathbb{R}$$

那么称X,Y相互独立

- $$X,Y\quad \text{相互独立}$$
$$\Updownarrow$$
$$P(X\in A,Y\in B)=P(X\in A)P(Y\in B),\quad \forall A,B\ \text{Borel 集}$$

- 如果X, Y相互独立，那么对任意Borel函数f, g，$f(X)$, $g(Y)$相互独立.

证明：由于f, g 是可测函数，那么对任意Borel 集A, B,
原像集$f^{-1}(A)$和$g^{-1}(B)$ 是Borel可测的。这样，

$$
\begin{aligned}
P(f(X) \in A, g(Y) \in B) &= P(X \in f^{-1}(A), Y \in g^{-1}(B)) \\
&\overset{\text{X,Y独立}}{=} P(X \in f^{-1}(A))P(Y \in g^{-1}(B)) \\
&= P(f(X) \in A)P(g(Y) \in B)
\end{aligned}
$$

因此，$f(X)$, $g(Y)$相互独立.

例. 如果X, Y相互独立，那么
(1) X^2, Y^2 相互独立;
(2) $|X|$, $|Y|$ 相互独立

- 多维随机向量

假设$(\Omega, \mathcal{A}, P)$是给定的概率空间，

$$\mathbf{X} = (X_1, X_2, \cdots, X_m) : \Omega \mapsto \mathbb{R}^m$$

如果对任意Borel 集$B \subset \mathbb{R}^m$，

$$\{\omega : \mathbf{X}(\omega) \in B\} \in \mathcal{A}$$

那么称$\mathbf{X}$为m-维随机向量

- m-元联合分布函数

假设$\mathbf{X} = (X_1, X_2, \cdots, X_m)$是$m$-维随机向量，其联合分布函数为

$$F_{\mathbf{X}}(\mathbf{x}) = P(X_1 \le x_1, \cdots, X_m \le x_m)$$

其中$\mathbf{x} = (x_1, \cdots, x_m)$.

- 边际分布

X_i的边际分布为

$$F_{X_i}(x_i) \quad = \quad P(X_i \le x_i)$$

- 独立随机变量

假设$\mathbf{X}=(X_1,X_2,\cdots,X_m)$是$m$-维随机向量，其联合分布函数为$F_{\mathbf{X}}(\mathbf{x})$，边际分布为$F_{X_i}(x_i)$,$i=1,\cdots,m$。如果

$$F_{\mathbf{X}}(\mathbf{x})=\prod_{i=1}^{m}F_{X_i}(x_i),\quad \forall\mathbf{x}=(x_1,\cdots,x_m)$$

那么称$X_1,X_2,\cdots,X_m$<u>相互独立</u>

- 注：如果$X_1,X_2,\cdots,X_m$相互独立，那么

(1) $X_{i_1},\cdots,X_{i_k}$, $k\le m$ 相互独立

(2) $f_1(X_1),f_2(X_2),\cdots,f_m(X_m)$相互独立

(3) $\mathbf{f}(x_1,x_2,\cdots,x_k)$, $\mathbf{g}(x_1,x_2,\cdots,x_l)$分别是$k$ 元和l 元Borel可测函数，其中$k+l\le m$。如果$A,B\subset 1,2,\cdots,m$，$\sharp A=k,\sharp B=l$，并且$A\cap B=\emptyset$，那么$\mathbf{f}(X_i,i\in A)$和$\mathbf{g}(X_i,i\in B)$相互独立

4. 随机向量的运算

- 随机向量的运算复杂多变，没有统一法则。

本节仅介绍一些常用的运算，如加、减、乘、除，线性变换和极值等

4.1 随机向量的加减

4.2 随机向量的乘除

4.3 随机向量的变换

4.4 极值随机变量

4.1 随机向量的加减

从随机变量的加、减开始

- 假设(X,Y)是离散型随机向量，分布为

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

求$Z=:X+Y$的分布

(1)Z的所有可能取值为x_i+y_j, $i,j=1,2,\cdots$。

将这些值重新记作$z_k, k=1,2,\cdots$，注意这些值互不相同；

(2) Z取每个值z_k的概率：

$$P(Z=z_k)=\sum_{i,j:x_i+y_j=z_k}p_{ij},\quad k=1,2,\cdots$$

特别，如果X,Y相互独立，那么可利用$p_{ij}=p_{i\cdot}p_{\cdot j}$，其中$p_{i\cdot}$是$X=x_i$的分布，$p_{\cdot j}$是$Y=y_j$的分布.

- 假设(X,Y)是连续型随机向量，具有分布密度$p(x,y)$

求$Z=:X+Y$的分布?

(1) Z仍是连续型随机变量，具有密度函数。需确定取值范围

(2) 计算Z的分布函数$F_Z(z)$

$$
\begin{aligned}
F_Z(z) &= P(X+Y\le z)\\
&= \int_{(x,y):x+y\le z} p(x,y)dxdy\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} p(x,y)dxdy\\
&= \int_{-\infty}^{z}\int_{-\infty}^{\infty} p(x,y-x)dxdy
\end{aligned}
$$

Z的密度函数$p_Z(z)$

$$p_Z(z)=\int_{-\infty}^{\infty} p(x,z-x)dx$$

- 当X,Y相互独立时，Z的密度函数$p_Z(z)$为

$$p_Z(z)=\int_{-\infty}^{\infty} p_X(x)p_Y(z-x)dx$$

- 减法类似

如，假设(X,Y)是连续型随机向量，具有分布密度$p(x,y)$

求$Z =: Y - X$的分布?

Z的密度函数$p_Z(z)$

$$p_Z(z) = \int_{-\infty}^{\infty} p(x, z+x)dx$$

当X,Y相互独立时，Z的密度函数$p_Z(z)$为

$$p_Z(z) = \int_{-\infty}^{\infty} p_X(x)p_Y(z+x)dx$$

- 例1. Bernoulli 随机变量之和

假设$X \sim B(n,p), Y \sim B(m,p)$，并且$X, Y$相互独立。求$X+Y$的分布?

(1) $X+Y$取值为$0, 1, 2\cdots, m+n$

(2) 取每个值的概率

$$
\begin{aligned}
P(X+Y=k) &= \sum_{l=0}^{k} P(X=l, Y=k-l) \\
&= \sum_{l=0}^{k} P(X=l)P(Y=k-l) \\
&= \sum_{l=0}^{k} \binom{n}{l} p^{l}(1-p)^{n-l} \binom{m}{k-l} p^{k-l}(1-p)^{m-(k-l)} \\
&= \binom{n+m}{k} p^{k}(1-p)^{m+n-k}
\end{aligned}
$$

- 独立Bernoulli 随机变量的和仍是Bernoulli 随机变量

- 例2. Poisson 随机变量之和

假设$X \sim \mathcal{P}(\lambda), Y \sim \mathcal{P}(\mu)$，并且$X, Y$相互独立。求$X+Y$的分布?

(1) $X+Y$的取值为$0, 1, 2, \cdots$

(2) 取每个值的概率

$$
\begin{aligned}
P(X+Y=k) &= \sum_{l=0}^{k} P(X=l, Y=k-l) \\
&= \sum_{l=0}^{k} P(X=l)P(Y=k-l) \\
&= \sum_{l=0}^{k} \frac{\lambda^l e^{-\lambda}}{l!} \cdot \frac{\mu^{k-l} e^{-\mu}}{(k-l)!} \\
&= \frac{(\lambda+\mu)^k}{k!} e^{-(\lambda+\mu)}
\end{aligned}
$$

- 独立Poisson 随机变量的和仍是Poisson 随机变量

例3. 均匀随机变量之和

假设$X \sim U(0,1), Y \sim U(0,1)$，并且$X, Y$相互独立。求$Z =: X+Y$的分布?

(1) Z的取值为$(0,2)$

(2) 取每个值的概率密度函数

当$0 < z \leq 1$

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} p_X(x) p_Y(z-x) dx \\ &= \int_0^z dx = z \end{aligned}$$

当$1 < z < 2$

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} p_X(x) p_Y(z-x) dx \\ &= \int_{z-1}^1 dx = 2 - z \end{aligned}$$

例4. 指数随机变量之和

假设$X \sim \exp(1), Y \sim \exp(1)$，并且$X, Y$相互独立。求$Z =: X + Y$的分布?

(1) Z的取值为$(0, \infty)$

(2) 取每个值的概率密度函数

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} p_X(x) p_Y(z-x) dx \\ &= \int_0^z e^{-x} e^{-(z-x)} dx \\ &= z e^{-z}, \quad z > 0 \end{aligned}$$

假设$X_i \sim \exp(1)$，并且$X_i, 1 \le i \le n$相互独立。求$Z =: \sum_{i=1}^n X_i$的分布?

$$p_Z(z) = \frac{1}{(n-1)!} z^{n-1} e^{-z}, \quad z > 0$$

- $\Gamma(\alpha, \beta)$ 分布

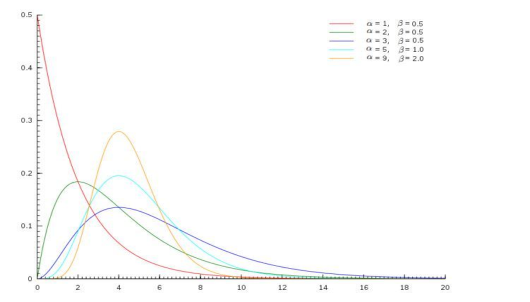
令$\alpha > 0$, $\beta > 0$ 定义

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

注意：$\Gamma(n) = (n-1)!$，$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.

$\Gamma(\alpha, \beta)$ 分布密度函数

$$p(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0$$



例5. 正态随机变量之和

假设$X\sim N(\mu_1,\sigma_1^2), Y\sim N(\mu_2,\sigma_2^2)$，并且$X,Y$相互独立。

求$Z =: X+Y$的分布?

(1) Z的取值为$(-\infty,\infty)$

(2) 取每个值的概率密度函数

$$
\begin{aligned}
p_Z(z) &= \int_{-\infty}^{\infty} p_X(x)p_Y(z-x)dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(z-x-\mu_2)^2}{2\sigma_2^2}} dx \\
&= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} e^{-\frac{(z-(\mu_1+\mu_2))^2}{2(\sigma_1^2+\sigma_2^2)}}
\end{aligned}
$$

- 独立正态随机变量的和仍是正态随机变量

4.2 随机向量的乘除

我们仅以连续型随机向量为例。离散型更简单些

假设(X,Y)是连续型随机向量，具有分布密度$p(x,y)$ 求$Z =: XY$的分布?

(1) Z 仍是连续型随机变量，需确定其取值范围

(2) 计算Z的分布函数$F_Z(z)$

注意$X,Y=0$的概率为0

$$\begin{aligned}
F_Z(z) &= P(XY \leq z) \\
&= \int_{(x,y):xy\leq z} p(x,y)dxdy \\
&= \int_{-\infty}^{0}\int_{z/x}^{\infty} p(x,y)dydx + \int_{0}^{\infty}\int_{-\infty}^{z/x} p(x,y)dydx \\
&= -\int_{-\infty}^{z}\int_{-\infty}^{0} \frac{1}{x}p(x,\frac{y}{x})dxdy + \int_{-\infty}^{z}\int_{0}^{\infty} \frac{1}{x}p(x,\frac{y}{x})dxdy
\end{aligned}$$

Z的密度函数$p_Z(z)$

$$\begin{aligned}
p_Z(z) &= -\int_{-\infty}^{0} \frac{1}{x}p(x,\frac{z}{x})dx + \int_{0}^{\infty} \frac{1}{x}p(x,\frac{z}{x})dx \\
&= \int_{-\infty}^{\infty} \frac{1}{|x|}p(x,\frac{z}{x})dx
\end{aligned}$$

- 当X,Y相互独立时，Z的密度函数$p_Z(z)$为

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{|x|} p(x, \frac{z}{x}) dx \\ &= \int_{-\infty}^{\infty} \frac{1}{|x|} p_X(x) p_Y(\frac{z}{x}) dx \end{aligned}$$

- 除法类似

假设(X,Y)是连续型随机向量，具有分布密度$p(x,y)$ 求$Z =: \frac{Y}{X}$的分布?

(1) Z 仍是连续型随机变量

(2) Z的密度函数$p_Z(z)$

$$p_Z(z) = \int_{-\infty}^{\infty} |x| p(x, zx) dx$$

- 当X, Y相互独立时，Z的密度函数$p_Z(z)$为

$$p_Z(z) = \int_{-\infty}^{\infty} |x| p_X(x) p_Y(zx) dx$$

例1. 假设X, Y相互独立，$X, Y \sim U(0,1)$。求$Z =: XY$的分布?

(1) Z的取值为$(0,1)$

(2) 取每个值的概率密度函数

$$
\begin{aligned}
p_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{|x|} p_X(x) p_Y(\frac{z}{x}) dx \\
&= \int_z^1 \frac{1}{x} p_X(x) p_Y(\frac{z}{x}) dx \\
&= \int_z^1 \frac{1}{x} dx \\
&= |\ln z|, \quad 0 < z < 1
\end{aligned}
$$

例2. 假设X, Y相互独立，$X, Y \sim U(0,1)$。求$Z =: \frac{Y}{X}$的分布?

(1) Z的取值为$(0,\infty)$; (2) 取每个值的概率密度函数

当$0 < z < 1$

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} |x| p_X(x) p_Y(zx) dx \\ &= \int_0^1 x dx = \frac{1}{2} \end{aligned}$$

当$z \geq 1$

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} |x| p_X(x) p_Y(zx) dx \\ &= \int_0^{\frac{1}{z}} x dx = \frac{1}{2z^2} \end{aligned}$$

即

$$p_Z(z) = \begin{cases} \frac{1}{2}, & 0 < z < 1 \\ \frac{1}{2z^2}, & z \geq 1 \end{cases}$$

例3. 假设$X \sim N(0,1), Y \sim N(0,1)$，并且$X, Y$相互独立。求$Z =: \frac{Y}{X}$的分布?

(1) Z的取值为$(-\infty, \infty)$

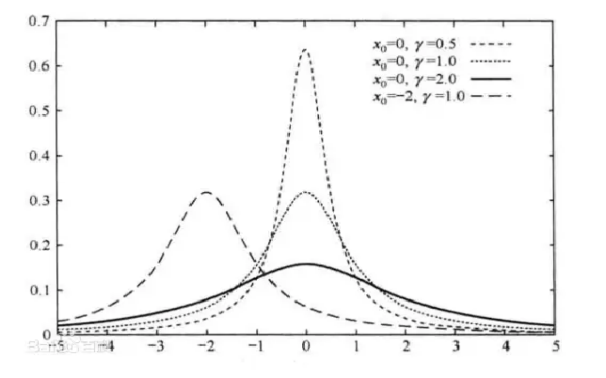
(2) 取每个值的概率密度函数

$$
\begin{aligned}
p_Z(z) &= \int_{-\infty}^{\infty} |x| p_X(x) p_Y(zx) dx \\
&= \int_{-\infty}^{\infty} |x| \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2x^2}{2}} dx \\
&= \frac{1}{\pi} \int_0^{\infty} x e^{-\frac{x^2(1+z^2)}{2}} dx \\
&= \frac{1}{\pi} \cdot \frac{1}{1+z^2}, \quad -\infty < z < \infty
\end{aligned}
$$

称$Z \sim p_Z(z)$为标准Cauchy 分布。

Cauchy 分布密度函数：

$$p(x;x_0,\gamma)=\frac{1}{\pi}\frac{\gamma}{\gamma^2+(x-x_0)^2},\quad x\in\mathbb{R}$$



4.3 随机向量的变换

- 我们仅考虑连续型随机向量的变换。

假设(X,Y)为连续型随机向量，具有联合密度函数$p_{(X,Y)}(x,y)$.
变换如下：

$$\begin{cases} U = f_1(X,Y) \\ V = f_2(X,Y) \end{cases}$$

求(U,V)的分布？

- 基本方法

$$\begin{aligned} & P(U \leq u, V \leq v) \\ & = P(f_1(X,Y) \leq u, f_2(X,Y) \leq v) \\ & = P((X,Y) \in \{(x,y) : f_1(x,y) \leq u, f_2(x,y) \leq v\}) \\ & = \int_{(x,y):f_1(x,y) \leq u, f_2(x,y) \leq v} p(x,y)dxdy \end{aligned}$$

一般情况下，很难给出一般性定理计算(U,V)的分布. 依赖于f_1, f_2所确定的积分区域。仅考虑特殊情形：假设f_1, f_2存在逆变换，即

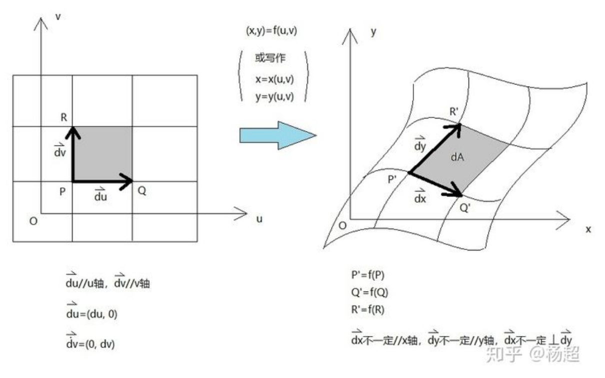
$$\begin{cases} X = g_1(U,V) \\ Y = g_2(U,V) \end{cases}$$

进一步假设g_1, g_2可导，Jacobi变换存在，行列式为

$$J = \det \begin{pmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{pmatrix}$$

那么(U,V)是连续型随机向量，具有联合密度函数$p_{(U,V)}(u,v)$。对于u,v属于f_1, f_2的值域，有

$$p_{(U,V)}(u,v) = p_{(X,Y}(g_1(u,v), g_2(u,v))|J|$$


v
u
O
R
P
Q
$\vec{dv}$
$\vec{du}$
(x,y)=f(u,v)
或写作
x=x(u,v)
y=y(u,v)
y
x
O
R'
P'
Q'
$\vec{dy}$
$\vec{dx}$
dA
$\vec{du}$//u轴，$\vec{dv}$//v轴
$\vec{du}$=(du, 0)
$\vec{dv}$=(0, dv)
P'=f(P)
Q'=f(Q)
R'=f(R)
$\vec{dx}$不一定//x轴，$\vec{dy}$不一定//y轴，$\vec{dx}$不一定⊥$\vec{dy}$
知乎 @杨超

$$\because \begin{bmatrix} \overrightarrow{dx} \\ \overrightarrow{dy} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \cdot \begin{bmatrix} \overrightarrow{du} \\ \overrightarrow{dv} \end{bmatrix}$$

$$\therefore \overrightarrow{dx} \times \overrightarrow{dy} = \left(\frac{\partial x}{\partial u} \cdot \overrightarrow{du} + \frac{\partial x}{\partial v} \cdot \overrightarrow{dv} \right) \times \left(\frac{\partial y}{\partial u} \cdot \overrightarrow{du} + \frac{\partial y}{av} \cdot \overrightarrow{dv} \right)$$

$$= \frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial v} \cdot \overrightarrow{du} \times \overrightarrow{dv} + \frac{\partial x}{\partial v} \cdot \frac{\partial y}{\partial u} \cdot \overrightarrow{dv} \times \overrightarrow{du}$$

$$= \frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial v} \cdot \overrightarrow{du} \times \overrightarrow{dv} - \frac{\partial x}{\partial v} \cdot \frac{\partial y}{\partial u} \cdot \overrightarrow{du} \times \overrightarrow{dv}$$

$$= \left(\frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \cdot \frac{\partial y}{\partial u} \right) \overrightarrow{du} \times \overrightarrow{dv}$$

$$= \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \overrightarrow{du} \times \overrightarrow{du}$$

例1. 假设$X \sim N(0,1), Y \sim N(0,1)$，并且$X, Y$相互独立。定义

$$\begin{cases} U = X + Y \\ V = X - Y \end{cases}$$

求(U, V)的分布密度？

逆变换：

$$\begin{cases} X = \frac{1}{2}(U+V) \\ Y = \frac{1}{2}(U-V) \end{cases}$$

Jacobi行列式为

$$J = \frac{1}{2}$$

(X, Y)的联合密度函数为

$$p_{X,Y}(x,y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

(U,V)的联合密度函数为

$$\begin{aligned} p_{U,V}(u,v) &= \frac{1}{2\pi}e^{-\frac{(\frac{1}{2}(u+v))^2+(\frac{1}{2}(u-v))^2}{2}}|J| \\ &= \frac{1}{4\pi}e^{-\frac{u^2+v^2}{4}} \end{aligned}$$

因此，U,V相互独立，且

$$U \sim N(0,2), \quad V \sim N(0,2)$$

例2. 假设$(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$。定义

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \end{pmatrix}$$

求(U,V)的分布密度？

记

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

并假设A可逆

$$\begin{pmatrix} X \\ Y \end{pmatrix} = A^{-1} \cdot \begin{pmatrix} U \\ V \end{pmatrix}$$

- Jacobi 行列式

$$J = |A^{-1}| = |A|^{-1}$$

- (X, Y)联合密度函数可以写成

$$p_{(X,Y)}(x,y) = \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})}$$

其中

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \vec{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

- (U,V)联合密度函数为

$$\begin{aligned} p_{(U,V)}(u,v) &= \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(A^{-1}\vec{u}-\vec{\mu})'\Sigma^{-1}(A^{-1}\vec{u}-\vec{\mu})} ||A|^{-1}| \\ &= \frac{1}{2\pi|A'\Sigma A|^{\frac{1}{2}}} e^{-\frac{1}{2}(\vec{u}-A\vec{\mu})'(A'\Sigma A)^{-1}(\vec{u}-A\vec{\mu})} \end{aligned}$$

其中

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} u \\ v \end{pmatrix}$$

因此，

$$(U,V) \sim N(A\vec{\mu}, A'\Sigma A)$$

将$A\vec{\mu}, A'\Sigma A$ 展开，可写出5个参数。

联合正态随机向量的线性组合仍是联合正态随机向量

例3. 假设$X \sim N(0,1), Y \sim N(0,1)$，定义

$$\begin{cases} \rho = \sqrt{X^2+Y^2} \\ \theta = \arctan \frac{Y}{X} \end{cases}$$

求(ρ,θ)的分布密度?

逆变换为

$$\begin{cases} X = \rho\cos\theta \\ Y = \rho\sin\theta \end{cases}$$

•Jacobi 行列式

$$J = \rho$$

- (ρ,θ)的联合密度函数为

$$\begin{aligned} p_{(\rho,\theta)}(\rho,\theta) &= \frac{1}{2\pi} e^{-\frac{1}{2}((\rho\cos\theta)^2+(\rho\sin\theta)^2)}\rho \\ &= \frac{1}{2\pi}\rho e^{-\frac{\rho^2}{2}} \end{aligned}$$

$$\rho \sim p_\rho(\rho) = \rho e^{-\frac{\rho^2}{2}}, \quad \rho > 0$$

$$\theta \sim p_\theta(\theta) = \frac{1}{2\pi}, \quad 0 < \theta < 2\pi$$

•称ρ为Rayleigh 分布，θ 为均匀分布

•ρ,θ相互独立

例4. 假设X, Y相互独立，服从参数为1的指数分布，即$X, Y \sim \exp(1)$. 求$U = X + Y, V = \frac{Y}{X}$的分布？

- 我们采用先求联合分布，再求边际分布的方法

注意

$$\begin{cases} U = X + Y \\ V = \frac{Y}{X} \end{cases}$$

逆变换为

$$\begin{cases} X = \frac{U}{1+V} \\ Y = \frac{UV}{1+V} \end{cases}$$

Jacobi 行列式为

$$J = \frac{U}{(1+V)^2}$$

(X,Y)的联合密度函数为

$$p_{(X,Y)}(x,y) = e^{-(x+y)}, \quad x,y>0$$

- (U,V)的联合密度函数为

$$\begin{aligned} p_{(U,V)}(u,v) &= e^{-(\frac{u}{1+v}+\frac{uv}{1+v})}\frac{u}{(1+v)^2} \\ &= ue^{-u}\cdot\frac{1}{(1+v)^2}, \quad u,v>0 \end{aligned}$$

由于上式中变量u,v分离，因此U,V相互独立. 容易看出

$$U \sim p_U(u) = ue^{-u}, \quad u>0$$

$$V \sim p_V(v) = \frac{1}{(1+v)^2}, \quad v>0$$

4.4 极值随机变量

假设$X_1, X_2, \cdots, X_n$是概率空间$(\Omega, \mathcal{A}, P)$上的随机变量。

对$\omega \in \Omega$，将$X_1(\omega), X_2(\omega), \cdots, X_n(\omega)$进行排序：

$$X_{(1)}(\omega) \leq X_{(2)}(\omega) \leq \cdots \leq X_{(n)}(\omega)$$

称$X_{(1)}, X_{(2)}, \cdots X_{(n)}$为次序随机变量

特别，$X_{(1)}$为极小值，$X_{(n)}$为极大值，$X_{(k)}$为第k小值

第一次考试ω_1：正常记录

$$X_1(\omega_1)=86, X_2(\omega_1)=73, X_3(\omega_1)=87, \cdots, X_{12}(\omega_1)=93$$

1	316010594	孙楠	德语	86	
2	316010518	王龙	工科试验班	73	
3	316010132	廖良	机科学与技术	87	
4	316010318	卢彪	机科学与技术	84	
5	316010384	黄诚	机科学与技术	97	
6	316010401	邱林	机科学与技术	89	
7	316010434	孙杰	工科试验班	89	
8	316010446	毛衡	工科试验班	93	
9	316010447	程坚	工科试验班	84	
10	316010437	程国	工科试验班	78	
11	316010452	吕橘	工科试验班	94	
12	316010452	薛悦	工科试验班	93	
13					

第一次考试ω_1：由低到高分记录

$$X_{(1)}(\omega_1) = X_2(\omega_1) = 73, X_{(2)}(\omega_1) = X_{10}(\omega_1) = 78$$

$$\cdots, X_{(12)}(\omega_1) = X_5(w_1) = 97$$

2	316010518	王龙	工科试验班	73	
10	316010437	程国	工科试验班	78	
4	316010318	卢彪	机科学与技术	84	
9	316010447	程坚	工科试验班	84	
1	316010594	孙楠	德语	86	
3	316010132	廖良	机科学与技术	87	
6	316010401	邱林	机科学与技术	89	
7	316010434	孙杰	工科试验班	89	
8	316010446	毛衡	工科试验班	93	
12	316010452	薛悦	工科试验班	93	
11	316010452	吕樯	工科试验班	94	
5	316010384	黄诚	机科学与技术	97	

第二次考试ω_2：正常记录

$$X_1(\omega_1) = 84, X_2(\omega_1) = 80, X_3(\omega_1) = 93, \cdots, X_{12}(\omega_1) = 92$$

1	316010594	孙楠	德语	84
2	316010518	王龙	工科试验班	80
3	316010132	廖良	机科学与技术	93
4	316010318	卢彪	机科学与技术	66
5	316010384	黄诚	机科学与技术	95
6	316010401	邱林	机科学与技术	86
7	316010434	孙杰	工科试验班	79
8	316010446	毛衡	工科试验班	89
9	316010447	程坚	工科试验班	88
10	316010437	程国	工科试验班	83
11	316010452	吕福	工科试验班	82
12	316010452	薛悦	工科试验班	92
13				

第二次考试ω_2：由低到高分记录

$$X_{(1)}(\omega_1) = X_4(\omega_1) = 66, X_{(2)}(\omega_1) = X_7(\omega_1) = 79$$

$$\cdots, X_{(12)}(\omega_1) = X_5(w_1) = 96$$

4	316010318	卢彪	机科学与技术	66
7	316010434	孙杰	工科试验班	79
2	316010518	王龙	工科试验班	80
11	316010452	吕榴	工科试验班	82
10	316010437	程国	工科试验班	83
1	316010594	孙楠	德语	84
6	316010401	邱林	机科学与技术	86
9	316010447	程坚	工科试验班	88
8	316010446	毛衡	工科试验班	89
12	316010452	薛悦	工科试验班	92
3	316010132	廖良	机科学与技术	93
5	316010384	黄诚	机科学与技术	95
13				

第三次考试ω_3：正常记录

$$X_1(\omega_1) = 85, X_2(\omega_1) = 96, X_3(\omega_1) = 98, \cdots, X_{12}(\omega_1) = 99$$

1	316010594	孙楠	德语	85
2	316010518	王龙	工科试验班	96
3	316010132	廖良	机科学与技术	98
4	316010318	卢彪	机科学与技术	93
5	316010384	黄诚	机科学与技术	91
6	316010401	邱林	机科学与技术	67
7	316010434	孙杰	工科试验班	83
8	316010446	毛衡	工科试验班	94
9	316010447	程坚	工科试验班	92
10	316010437	程国	工科试验班	99
11	316010452	吕儒	工科试验班	81
12	316010452	薛悦	工科试验班	99
13				

第三次考试ω_3：由低到高分记录

$$X_{(1)}(\omega_1) = X_6(\omega_1) = 67, X_{(2)}(\omega_1) = X_{11}(\omega_1) = 81$$

$$\cdots, X_{(12)}(\omega_1) = X_{12}(w_1) = 99$$

6	316010401	邱林	机科学与技术	67
11	316010452	吕榴	工科试验班	81
7	316010434	孙杰	工科试验班	83
1	316010594	孙楠	德语	85
5	316010384	黄诚	机科学与技术	91
9	316010447	程坚	工科试验班	92
4	316010318	卢彪	机科学与技术	93
8	316010446	毛衡	工科试验班	94
2	316010518	王龙	工科试验班	96
3	316010132	廖良	机科学与技术	98
10	316010437	程国	工科试验班	99
12	316010452	薛悦	工科试验班	99
13				

- 极大值

假设$X_1, X_2, \cdots, X_n$是独立随机变量，具有相同分布函数$F(x)$。求$X_{(n)}$的分布？

$$\begin{aligned} F_{X_{(n)}}(x) &= P(X_{(n)} \leq x) = P(\max_{1\leq i\leq n} X_i \leq x) \\ &= P(\bigcap_{i=1}^{n}\{X_i \leq x\}) \\ &= \prod_{i=1}^{n} P(X_i \leq x) \\ &= F^n(x) \end{aligned}$$

- 假设分布函数$F(x)$有密度函数，那么$X_{(n)}$ 也具有密度函数

$$p_{X_{(n)}}(x) = nF(x)^{n-1}F'(x)$$

- 极小值

假设$X_1, X_2, \cdots, X_n$是独立随机变量，具有相同分布函数$F(x)$。求$X_{(1)}$的分布？

$$\begin{aligned} P(X_{(1)} > x) &= P(\min_{1\le i\le n} X_i > x) \\ &= P(\bigcap_{i=1}^{n}\{X_i > x\}) = \prod_{i=1}^{n} P(X_i > x) = (1 - F(x))^n \end{aligned}$$

分布函数为

$$\begin{aligned} F_{X_{(1)}}(x) &= P(X_{(1)} \le x) = 1 - P(X_{(1)} > x) \\ &= 1 - \prod_{i=1}^{n}(1 - F(x))^n \end{aligned}$$

- 假设分布函数$F(x)$有密度函数，那么$X_{(1)}$ 也具有密度函数

$$p_{X_{(1)}}(x) = n(1 - F(x))^{n-1} F'(x)$$

- 第k 小值

假设$X_1, X_2, \cdots, X_n$是独立连续随机变量，具有相同分布函数$F(x)$, 密度函数$p(x)$。求$X_{(k)}$的分布密度？

$X_{(k)}$的分布密度为

$$p_{X_{(k)}}(x) = (n-k+1)\binom{n}{k-1}F^{k-1}(x)p(x)(1-F(x))^{n-k}$$

例. 假设$X_1, X_2, \cdots, X_n$是独立连续随机变量，服从$(0,1)$上均匀分布。求$X_{(1)}$，$X_{(n)}$的分布密度？

$X_{(1)}$的分布密度

$$p_{X_{(1)}}(x) = n(1-x)^{n-1}, \quad 0 < x < 1$$

$X_{(n)}$的分布密度

$$p_{X_{(n)}}(x) = nx^{n-1}, \quad 0 < x < 1$$