- 随机变量— Random Variable

In the late 1940s, William Feller (1906-1970) 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.

- 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

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,

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

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

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

- 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

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

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

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。

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

(2) 二项展开系数

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

(3) X最可能的值

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

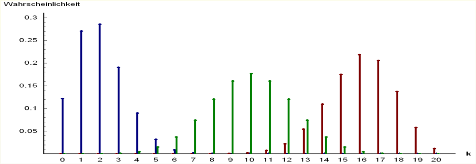
所以，

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 ≤ x)	
Cumulative probability: P(X > x)	
Cumulative probability: P(X ≥ 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)

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

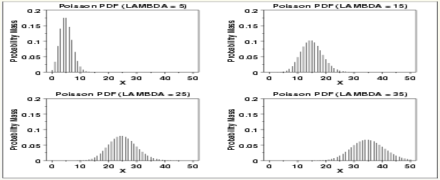
(2) Poisson分布是Poisson过程的基础，用于描述随机服务系统

(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 exactly once; at least once?

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

$$p = \frac{4}{2598900} = \frac{1}{649740}$$

Let $\lambda = np$. 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

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，$1,2,\cdots$

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

- 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$n(\leq N)$件，用X表示n件产品中次品数。

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

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

- 证明

$$\sum_{k=0}^{\min\{M,n\}} \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = 1$$

- 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；类似于物体质量密度。
- 注：

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

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

- 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

(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

(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

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

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：

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

$$\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dxdy &= \int_{0}^{\infty}\int_{0}^{2\pi} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r dr d\theta \\ &= 1 \end{aligned}$$

- 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

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

(10) 正态分布表可查

- 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，函数

$$X : \Omega \mapsto R$$

如果对任意Borel集B，

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

那么称X为随机变量

特别，

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

- 分布函数

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

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

称F为X的分布函数。

- 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

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

$F(x)$ 是一个阶梯型函数

例. 假定

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

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

- 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

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

$F(x)$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。
- $\Phi(x)$的值可以通过查表得到

如，

$$\Phi(1) =, \quad \Phi(2) =, \quad \Phi(3) =$$

- 分布函数的性质

回忆

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

F具有下列性质

(1)

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

(2) $F(x)$ 单调不减

(3) $F(x)$ 左极限存在，右连续

上述性质很容易从概率P的性质推导

- 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$X : \Omega \mapsto R$，使得

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

- 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

(2) 如何收取保费依赖于寿命长短，并不是简单的比例关系

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$X \sim \exp(\lambda)$，

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

求$Y = f(X)$的分布？

对任意$-\infty < y < \infty$，

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

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$-\infty < y < \infty$，

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

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

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. 假设X具有连续的分布函数$F(x)$，求$Y=F(X)$的分布？

注意，$F(x)$具有良好的性质：
$0\le F(x)\le 1$，单调不减，右连续.
定义
$$F^{-1}(y)=\inf\{x:F(x)>y\}$$

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

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

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

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

- $Y \sim F(x)$

- 人们经常需要研究和使用随机向量

(1) 抽查高考考生成绩

(2) 婴儿体检

(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

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

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

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

- 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. 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

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

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

- (2)无放回

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

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

- 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，那么我们可以得到条件分布:

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，

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

- 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. 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

- 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，X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

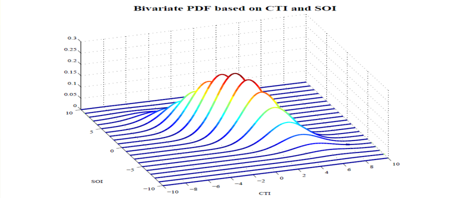
如果随机向量(X,Y)具有密度函数: $\forall x, y \in R$

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

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

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


Bivariate PDF based on CTI and SOI
0.3
0.25
0.2
0.15
0.1
0.05
0
10
5
0
−5
−10
SOI
−10
−8
−6
−4
−2
0
2
4
6
8
10
CTI

•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

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

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

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

- 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？

$$X\sim p_X(x)$$

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

进行配方得，

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

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$P(Y\leq y|X=x)$

注意$P(X=x)=0$，我们采用

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

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

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$p_{X|Y}(x|y),p_{Y|X}(y|x)$?

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

进行整理得，

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

即固定y,

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

类似地，固定x,

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

- 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

- 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

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

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

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

- $F(x,y)$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)

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

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

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

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

- 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

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

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

如，如果X, Y相互独立，那么X^2, Y^2也相互独立

例.

- 多维随机向量

假设$(\Omega, \mathcal{A}, P)$是给定的概率空间，

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

如果对任意Borel 集$B \subset R^m$，

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

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

$$\begin{aligned} F_{X_i}(x_i) &= P(X_i \le x_i) \\ &= F(+\infty, \cdots, x_i, \cdots, +\infty) \end{aligned}$$

- 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$F_{X_i}(x_i)$,$i = 1, \cdots, m$。如果对所有$\mathbf{x} = (x_1, \cdots, x_m)$，

$$F_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{m} F_{X_i}(x_i)$$

那么称$X_1, X_2, \cdots, X_m$相互独立

- 注：如果$X_1, X_2, \cdots, X_m$相互独立，那么

(1) $X_{i_1}, \cdots, X_{i_k}$, $k \leq m$ 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$k + l \leq m$。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

- 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

- 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$x_i + y_j,\ i, j = 1, 2, \cdots$。
将这些值重新记做$z_k, k = 1, 2, \cdots$，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

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

$$
\begin{aligned}
&= \sum_{l=0}^{k} C_n^l p^l (1-p)^{n-l} C_m^{k-l} p^{k-l} (1-p)^{m-(k-l)} \\
&= C_{n+m}^k p^k (1-p)^{m+n-k}
\end{aligned}
$$

- 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

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

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

- 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

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

- 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

- 随机向量的乘、除运算

我们仅以连续型随机向量为例。离散型更简单些

m 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$X \sim U(0,1), Y \sim U(0,1)$，并且$X, Y$相互独立。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$X \sim U(0,1), Y \sim U(0,1)$，并且$X, Y$相互独立。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

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

- 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

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

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

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

求(r,θ)的分布密度？

逆变换为

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

- Jacobi 行列式

$$J = r$$

- (r,θ)的联合密度函数为

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

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

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

- 称r为Rayleigh 分布，θ 为均匀分布
- r, 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

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

由于上式中变量u, v分离，因此U, V相互独立。

容易看出

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

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

- 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

- 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$p_{X_{(n)}}(x)$，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？

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

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$p_{X_{(1)}}(x)$，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

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

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

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