第二章　随机变量与分布函数

Tianxiao Pang

Zhejiang University

September 12, 2021

内容

内容

内容

内容

内容

内容

离散型随机变量及其分布

2.1.1 随机变量(random variable)的概念

离散型随机变量及其分布

2.1.1 随机变量(random variable)的概念

随机试验有多种类型的结果, 如数值型结果、非数值型结果. 问题: 如何能方便地(用现代数学的工具)把这些结果及其相应的概率大小表达出来?

离散型随机变量及其分布

2.1.1 随机变量(random variable)的概念

随机试验有多种类型的结果, 如数值型结果、非数值型结果. 问题: 如何能方便地(用现代数学的工具)把这些结果及其相应的概率大小表达出来?

答案: 随机变量. (定义在样本空间上的函数, 能把样本空间里的元素映射到实数轴, 以便用现代数学的方法来研究概率问题)

例如:

(1) 对某物体的长度进行测量, 一切可能的测量值构成一个样本空间$\Omega = \{\omega | \omega \in (a, b)\}$. 那么可用一变量$\xi$跟测量的结果联系起来: $\xi(\omega) = \omega$. "$\xi \in [1.5, 2.5]$" 表示事件{测量值在1.5与2.5之间}.

例如:

(1) 对某物体的长度进行测量, 一切可能的测量值构成一个样本空间$\Omega = \{\omega | \omega \in (a, b)\}$. 那么可用一变量$\xi$跟测量的结果联系起来: $\xi(\omega) = \omega$. "$\xi \in [1.5, 2.5]$"表示事件{测量值在1.5与2.5之间}.

(2) 抛一枚硬币, 观察正反面. 一切可能的结果构成一个样本空间$\Omega = \{\omega_1, \omega_2\} = \{正面, 反面\}$. 那么可令

$$\xi(\omega_1) = 1, \ \xi(\omega_2) = 0.$$

由于随机变量本身取实数值，所以我们通常感兴趣的是$\xi(\omega)$取值于单点集、区间$(a,b]$，或若干个这种区间的并、交. 因此我们希望$\{\xi(\omega)\in A\}$ (其中A是由左开右闭区间经并、交、逆等运算而得到的直线上的某一个点集，即一维Borel集)有概率可言.

在概率的公理化定义里我们只对概率空间$(\Omega,\mathscr{F},\mathsf{P})$的事件域$\mathscr{F}$中的事件才定义了概率，因此我们自然要求$\{\omega:\xi(\omega)\in A\}$(其中$A$ 是一维Borel集)属于$\mathscr{F}$.

定义（随机变量）

设$\xi(\omega)$是定义在概率空间$(\Omega, \mathscr{F}, P)$上的单值实函数，且对于任一Borel集B，有

$$\xi^{-1}(B) = \{\omega : \xi(\omega) \in B\} \in \mathscr{F}, \tag{2.1.1}$$

就称$\xi(\omega)$为随机变量(random variable)，而称$P(\{\omega : \xi(\omega) \in B\})$，$B \in \mathscr{B}$(一维Borel σ域)，为随机变量$\xi(\omega)$的概率分布(probability distribution).

注: (1) 通常用希腊字母$\xi, \eta, \zeta, \cdots$ 或大写英文字母$X, Y, Z, \cdots$ 等等来表示随机变量;

(2) 完整地写出每个随机变量的概率分布通常是不容易做到的, 因为概率分布是对随机变量的完整描述. 但是在很多情况下, 我们只需对一系列特殊的Borel集求得(2.1.1)的概率就能决定整个概率分布了;

(3) 今后, 我们通常把$\xi(\omega)$简写称ξ, $\{\omega : \xi(\omega) \in B\}$简写成$\{\xi(\omega) \in B\}$. 因此

$$\mathsf{P}(\xi \in B) = \mathsf{P}(\xi(\omega) \in B) = \mathsf{P}(\{\omega : \xi(\omega) \in B\}).$$

2.1.2 离散型随机变量

2.1.2 离散型随机变量

定义（离散型随机变量）

若随机变量ξ可能取的值至多可列个(有限个或可列无穷多个), 则称ξ为离散型随机变量(discrete random variables).

对于离散型随机变量ξ, 我们往往关心两方面的内容: (1) 可能取值的集合$\{x_i; i=1,2,\cdots\}$; (2) ξ取这些值的概率, 即

$$p_i := p(x_i) := \mathsf{P}(\xi = x_i), i = 1, 2, \cdots .$$

我们称

对于离散型随机变量ξ, 我们往往关心两方面的内容: (1) 可能取值的集合$\{x_i; i=1,2,\cdots\}$; (2) ξ取这些值的概率, 即

$$p_i := p(x_i) := \mathsf{P}(\xi = x_i), i = 1, 2, \cdots .$$

我们称

$$\mathsf{P}(\xi = x_i) = p(x_i), i = 1, 2, \cdots$$

或

ξ	x_1	x_2	$\cdots$	x_n	$\cdots$
P	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$\cdots$

为ξ的**分布列**(distribution sequence), 有时也称为ξ的概率分布.

分布列性质:

分布列性质:

- $p(x_i) \geq 0, i = 1, 2, \cdots;$
- $\sum_{i=1}^{\infty} p(x_i) = 1.$

分布列性质:

- $p(x_i) \geq 0, i = 1, 2, \cdots;$
- $\sum_{i=1}^{\infty} p(x_i) = 1.$

注: 有了分布列, 就可以求得与离散型随机变量ξ有关的一切事件的概率. 事实上, 由概率的可列可加性, 对直线上任一Borel集B, 有

$$\mathsf{P}(\xi(\omega) \in B) = \sum_{x_i \in B} p(x_i).$$

例1 设随机变量ξ的分布列为

$$\mathsf{P}(\xi = k) = \frac{c\lambda^k}{k!}, \quad k = 0, 1, 2, \cdots, \quad \lambda > 0.$$

求常数c的值.

例1 设随机变量ξ的分布列为

$$\mathsf{P}(\xi = k) = \frac{c\lambda^k}{k!}, \quad k = 0, 1, 2, \cdots, \quad \lambda > 0.$$

求常数c的值.

解: 由分布列性质可知

$$1 = \sum_{k=1}^{\infty} p(x_k) = \sum_{k=0}^{\infty} \frac{c\lambda^k}{k!} = ce^{\lambda},$$

所以

$$c = e^{-\lambda}.$$

例2 在Bernoulli概型中, 每次成功的概率为p. 记直至得到第r次成功时的试验次数为ξ, 求ξ的分布列.

例2 在Bernoulli概型中，每次成功的概率为p. 记直至得到第r次成功时的试验次数为ξ，求ξ的分布列.

解: ξ的取值情况: $\{r,\ r+1,\ r+2,\ \cdots\}$. 再来确定$\xi$取这些值的概率大小. 因为

$$\{\xi=k\}=\{\text{前}k-1\text{次试验中有}r-1\text{次成功},\ k-r\text{次不成功且第}k\text{次试验成功}\}.$$

所以

$$\begin{aligned}\mathrm{P}(\xi=k)&=\binom{k-1}{r-1}p^{r-1}q^{k-r}p\\&=\binom{k-1}{r-1}p^{r}q^{k-r},\quad k=r,r+1,r+2,\cdots,\qquad(2.1.2)\end{aligned}$$

其中$q=1-p$. (2.1.2)式即为ξ的分布列，通常被称为巴斯卡分布(Pascal distribution).

一些常见的、重要的离散型随机变量:

一些常见的、重要的离散型随机变量:

1. 退化分布(degenerate distribution)

一些常见的、重要的离散型随机变量:

1. 退化分布(degenerate distribution)

定义

设随机变量ξ只取一个常数值c, 即

$$P(\xi = c) = 1,$$

我们称它为退化分布, 又称为单点分布.

2. 两点分布(two-point distribution)

2. 两点分布(two-point distribution)

定义

若一个随机变量只取两个值x_1, x_2, 且相应的分布列为

ξ	x_1	x_2
P	p	q

其中$p, q > 0, q = 1 - p$. 则称ξ服从两点分布.

一个特殊的两点分布: Bernoulli分布(也称$0-1(p)$分布)

一个特殊的两点分布: Bernoulli分布(也称$0-1(p)$分布)

在Bernoulli试验中, 每次试验只有两个结果: $A, \bar{A}$. 我们可用一个随机变量(这里, 通常采用A的示性函数)把这两个试验结果数量化. 即令

一个特殊的两点分布: Bernoulli分布(也称$0-1(p)$分布)

在Bernoulli试验中, 每次试验只有两个结果: $A, \bar{A}$. 我们可用一个随机变量(这里, 通常采用A的示性函数)把这两个试验结果数量化. 即令

$$\xi = I\{A\} = \begin{cases} 1, & \text{如果}A\text{发生;} \\ 0, & \text{如果}A\text{不发生.} \end{cases}$$

其分布列为

一个特殊的两点分布: Bernoulli分布(也称$0-1(p)$分布)

在Bernoulli试验中, 每次试验只有两个结果: $A, \bar{A}$. 我们可用一个随机变量(这里, 通常采用A的示性函数)把这两个试验结果数量化. 即令

$$\xi = I\{A\} = \begin{cases} 1, & \text{如果}A\text{发生;} \\ 0, & \text{如果}A\text{不发生.} \end{cases}$$

其分布列为

ξ	0	1
P	q	p

此时, 可记$\xi \sim 0-1(p)$.

任一个Bernoulli试验的结果(只有两个), 例如抛硬币得“正面”与“反面”, 产品“合格”与“不合格”, $\cdots$等等, 都可以用Bernoulli分布描述.

3. 二项分布(binomial distribution)

3. 二项分布(binomial distribution)

定义

若随机变量ξ的分布列为

$$P(\xi = k) = \binom{n}{k} p^k q^{n-k} =: b(k; n, p), \quad k = 0, 1, 2, \cdots, n,$$

其中$p, q > 0, p + q = 1$. 则称ξ服从参数为n, p的二项分布，记作$\xi \sim B(n, p)$.

$n = 1$时的二项分布即为$0 - 1(p)$分布.

二项分布的应用很广, 例如:

二项分布的应用很广，例如:

(1) 保险公司对某种灾害(汽车被盗，火灾等)保险，各人发生此种灾害与否可认为相互独立，并假定概率相等. 设一年间一人发生此种灾害的概率为p，则在参加此种保险的n人中发生此种灾害的人数η服从二项分布.

二项分布的应用很广, 例如:

(1) 保险公司对某种灾害(汽车被盗, 火灾等)保险, 各人发生此种灾害与否可认为相互独立, 并假定概率相等. 设一年间一人发生此种灾害的概率为p, 则在参加此种保险的n人中发生此种灾害的人数η服从二项分布.

(2) n台同类机器, 在一段时间内每台损坏的概率为p, 则在这段时间内损坏的机器总数服从二项分布.

二项分布的性质:

二项分布的性质:

(1) $b(k; n, p) = b(n-k; n, 1-p)$.

二项分布的性质:

(1) $b(k;n,p)=b(n-k;n,1-p)$. 这是因为$\binom{n}{k}=\binom{n}{n-k}$, 或理解成: 在Bernoulli概型中, $\{k$次成功$\}=\{n-k$次不成功$\}$.

二项分布的性质:

(1) $b(k;n,p)=b(n-k;n,1-p)$. 这是因为$\binom{n}{k}=\binom{n}{n-k}$, 或理解成: 在Bernoulli概型中, $\{k$次成功$\}=\{n-k$次不成功$\}$.

(2) 单调增减性以及最可能成功次数.

固定n,p. 由于

$$\frac{b(k;n,p)}{b(k-1;n,p)}=$$

二项分布的性质:

(1) $b(k;n,p)=b(n-k;n,1-p)$. 这是因为$\binom{n}{k}=\binom{n}{n-k}$, 或理解成: 在Bernoulli概型中, $\{k$次成功$\}=\{n-k$次不成功$\}$.

(2) 单调增减性以及最可能成功次数.

固定n,p. 由于

$$\frac{b(k;n,p)}{b(k-1;n,p)}=\frac{\binom{n}{k}p^k(1-p)^{n-k}}{\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1}}=1+\frac{(n+1)p-k}{k(1-p)},$$

因此

当$k<(n+1)p$时, $\frac{b(k;n,p)}{b(k-1;n,p)}>1$, $b(k;n,p)$单调增加;
当$k>(n+1)p$时, $\frac{b(k;n,p)}{b(k-1;n,p)}<1$, $b(k;n,p)$单调减少.

- 当$(n+1)p$是整数时,

$$b(k;n,p)=b(k-1;n,p),\quad k=(n+1)p.$$

$k=(n+1)p$或$k=(n+1)p-1$为最可能成功次数.

- 当$(n+1)p$是整数时,

$$b(k;n,p)=b(k-1;n,p),\quad k=(n+1)p.$$

$k=(n+1)p$或$k=(n+1)p-1$为最可能成功次数.

- 当$(n+1)p$不是整数时，由上面的单调性分析以及

$$b([(n+1)p]+1;n,p)<b([(n+1)p];n,p),$$

可知$[(n+1)p]$为最可能成功次数(这里，$[\cdot]$表示取整符号).

(3) 递推公式.

(3) 递推公式.

设$\xi \sim B(n,p)$, 则

$$\mathsf{P}(\xi = k+1) = \frac{p(n-k)}{q(k+1)}\mathsf{P}(\xi = k).$$

此公式容易由二项分布的表达式得到. 从$\mathsf{P}(\xi = 0) = q^n$出发, 人们可以用此公式递推得到各个$\mathsf{P}(\xi = k)$的值.

(3) 递推公式.

设$\xi \sim B(n,p)$, 则

$$\mathsf{P}(\xi = k+1) = \frac{p(n-k)}{q(k+1)}\mathsf{P}(\xi = k).$$

此公式容易由二项分布的表达式得到. 从$\mathsf{P}(\xi = 0) = q^n$出发, 人们可以用此公式递推得到各个$\mathsf{P}(\xi = k)$的值.

此外, 我们也可以使用各种数学和统计软件, 如Matlab、SPSS、R、Python等, 计算概率分布值.

(4) $n \to \infty$时的渐近性质.

(4) $n \to \infty$时的渐近性质.

假定p与n有关, 记作p_n. 我们有下列的“二项分布的泊松(Poisson)逼近”.

(4) $n \to \infty$时的渐近性质.

假定p与n有关, 记作p_n. 我们有下列的“二项分布的泊松(Poisson)逼近”.

定理 (Poisson定理)

如果存在正常数λ, 当$n \to \infty$时, 有$np_n \to \lambda$. 则

$$\lim_{n\to\infty} b(k;n,p_n) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0, 1, 2, \cdots .$$

证明: 利用一个初等不等式

$$|a^n - b^n| \leq n|a-b|, \text{ 如果}|a| \leq 1, |b| \leq 1.$$

我们可得, 当$n \to \infty$时,

$$\left|(1-p_n)^n - (1-\frac{\lambda}{n})^n\right| \leq n\left|p_n - \frac{\lambda}{n}\right| \to 0.$$

证明: 利用一个初等不等式

$$|a^n - b^n| \le n|a-b|, \text{ 如果}|a| \le 1, |b| \le 1.$$

我们可得, 当$n \to \infty$时,

$$\left|(1-p_n)^n - (1-\frac{\lambda}{n})^n\right| \le n\left|p_n - \frac{\lambda}{n}\right| \to 0.$$

记$\lambda_n = np_n$, 则$p_n = \lambda_n/n$. 固定k, 可得

$$\begin{aligned} b(k;n,p_n) &= \binom{n}{k} p_n^k (1-p_n)^{n-k} \\ =& \frac{n(n-1)\cdots(n-k-1)}{k!}\left(\frac{\lambda_n}{n}\right)^k (1-\frac{\lambda_n}{n})^{n-k} \\ =& \frac{\lambda_n^k}{k!} \cdot \frac{n(n-1)\cdots(n-k-1)}{n^k} \cdot (1-\frac{\lambda_n}{n})^n \cdot (1-\frac{\lambda_n}{n})^{-k} \\ \to& \frac{\lambda^k}{k!} e^{-\lambda} \quad (n \to \infty). \end{aligned}$$

注: 通常, p与n无关. 但实践表明: 当n很大($n \geq 50$), p很小($p \leq 0.1$), 而np不很大时, 可近似取$np = \lambda$, 且

$$b(k; n, p) \approx \frac{\lambda^k}{k!} e^{-\lambda}.$$

$\frac{\lambda^k}{k!} e^{-\lambda}$的计算要比$b(k; n, p)$的计算容易得多.

例3 某人每次射击时击中目标的概率为0.001. 射击5000次, 求击中两弹或两弹以上的概率.

例3 某人每次射击时击中目标的概率为0.001. 射击5000次, 求击中两弹或两弹以上的概率.

解: 设ξ为击中的弹数, 则$\xi \sim B(5000, 0.001)$. 所求概率为

$$\mathsf{P}(\xi \geq 2) = 1 - b(0; 5000, 0.001) - b(1; 5000, 0.001).$$

例3 某人每次射击时击中目标的概率为0.001. 射击5000次, 求击中两弹或两弹以上的概率.

解: 设ξ为击中的弹数, 则$\xi \sim B(5000, 0.001)$. 所求概率为

$$\mathsf{P}(\xi \geq 2) = 1 - b(0; 5000, 0.001) - b(1; 5000, 0.001).$$

因为这里n很大p很小, 且$np = 5$. 所以我们可利用Poisson定理做近似计算:

$$\mathsf{P}(\xi \geq 2) \approx 1 - \frac{5^0 e^{-5}}{0!} - \frac{5^1 e^{-5}}{1!} \approx 0.9596.$$

*Poisson定理给出当n很大, p很小时, 二项分布值的近似计算公式, 下面定理给出了当n很大, p大小适中时, 二项分布的近似.

*Poisson定理给出当n很大, p很小时, 二项分布值的近似计算公式, 下面定理给出了当n很大, p大小适中时, 二项分布的近似.

定理 (de Moivre-Laplace)

设$\xi_n \sim B(n,p)$, $p=p_n$, $q=1-p$满足$npq\to\infty$. 记

$$j=j(n), x=x(n)=\frac{j-np}{\sqrt{npq}}.$$

则在任何有限区间$[a,b]$上, 对$x\in[a,b]$一致地有

$$P_n(x)=P(\xi_n=j)\sim\frac{1}{\sqrt{2\pi npq}}e^{-x^2/2}.$$

其中j随n变化使得$x=x(n)$保持在有限区间$[a,b]$中, $a_n\sim b_n$表示$a_n/b_n\to 1$.

*de Moivre-Laplace定理告诉我们: 当npq很大, 而$\frac{j-np}{\sqrt{npq}}$不是很大时,

$$\mathrm{P}(\xi_n = j) \approx \frac{1}{\sqrt{2\pi npq}} e^{-x^2/2}, x = \frac{j-np}{\sqrt{npq}}.$$

进一步, 在de Moivre-Laplace定理的条件下还有如下积分形式的结果

$$\mathrm{P}\left(a \leq \frac{\xi_n - np}{\sqrt{npq}} \leq b\right) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} \mathrm{d}x. \tag{2.1.3}$$

这就是de Moivre-Laplace中心极限定理, 是第四章要介绍的一般中心极限定理的一个特例.

4. 泊松分布(Poisson distribution)

4. 泊松分布(Poisson distribution)

定义

若随机变量ξ的分布列为

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

则称ξ服从参数为λ的泊松分布，记为$\xi \sim P(\lambda)$.

4. 泊松分布(Poisson distribution)

定义

若随机变量ξ的分布列为

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

则称ξ服从参数为λ的泊松分布，记为$\xi \sim P(\lambda)$.

注: 在第三章，我们将知道λ的含义：ξ的平均值.

Poisson分布的应用:

Poisson分布的应用:

- 二项分布的近似计算(见Poisson定理).
- 可以用来描述离散型随机现象: 如果n个独立事件$A_1, \cdots, A_n$中每个发生的概率p很小, 那么这n个事件发生的次数近似服从Poisson分布$P(np)$. 因此, 通常我们认为: 单位时间/区间/面积/体积/$\cdots$ 中的计数过程服从Poisson分布. 例如, 一定时间内接到的电话呼叫数、某公交站台在一定时间内的上车人数、一本书的错别字个数、一平方米玻璃上的气泡数等等, 均可认为服从Poisson分布.

例4 考察通过某交叉路口的汽车流. 若在1分钟内没有汽车通过的概率为0.2, 求1分钟内有多于一辆汽车通过的概率.

例4 考察通过某交叉路口的汽车流. 若在1分钟内没有汽车通过的概率为0.2, 求1分钟内有多于一辆汽车通过的概率.

解: 设ξ为1分钟内通过的车辆数, 可假设$\xi \sim P(\lambda)$. 先来确定参数λ. 由题意,

$$\mathsf{P}(\xi = 0) = e^{-\lambda} = 0.2,$$

所以可得$\lambda = \ln 5$. 因此, 所求概率为

$$\begin{aligned}\mathsf{P}(\xi > 1) =& 1 - \mathsf{P}(\xi = 0) - \mathsf{P}(\xi = 1) \\ =& 1 - e^{-\lambda} - \lambda e^{-\lambda} \\ =& 1 - \frac{1}{5} - \frac{1}{5}\ln 5 \\ \approx& 0.4782.\end{aligned}$$

在Poisson近似中，事件$A_1, \cdots, A_n$的独立性要求可以放宽．例如，在第一章的匹配问题中，如果用A_i表示第i个士兵拿对了自己的枪这一事件，那么$\mathsf{P}(A_i) = \frac{1}{n}$．但是$A_1, \cdots, A_n$不相互独立，然而

$$\mathsf{P}(A_i|A_j) = \frac{1}{n-1} \approx \mathsf{P}(A_i), \quad i \neq j.$$

因此我们可以认为$A_1, \cdots, A_n$近似独立，从而认为n个士兵中拿对了自己的枪的人数ξ近似服从Poisson分布$P(n \times \frac{1}{n}) = P(1)$．所以没有士兵拿对自己的枪的概率为

在Poisson近似中，事件$A_1, \cdots, A_n$的独立性要求可以放宽．例如，在第一章的匹配问题中，如果用A_i表示第i个士兵拿对了自己的枪这一事件，那么$\mathsf{P}(A_i) = \frac{1}{n}$．但是$A_1, \cdots, A_n$不相互独立，然而

$$\mathsf{P}(A_i|A_j) = \frac{1}{n-1} \approx \mathsf{P}(A_i), \quad i \neq j.$$

因此我们可以认为$A_1, \cdots, A_n$近似独立，从而认为n个士兵中拿对了自己的枪的人数ξ近似服从Poisson分布$P(n \times \frac{1}{n}) = P(1)$．所以没有士兵拿对自己的枪的概率为

$$\mathsf{P}(\xi = 0) \approx \frac{1^0}{0!} e^{-1} = e^{-1},$$

从而至少有一个士兵拿对自己的枪的概率近似等于$1 - e^{-1}$，与第一章结论相符．

一般地, 设有n个事件$A_1, \cdots, A_n$, 第i个事件发生的概率为p_i. 如果这些概率p_i都很小, 而且这些事件相互独立或者近似独立, 那么这些事件发生的次数近似服从Poisson分布$P(\lambda)$, 其中

$$\lambda = \sum_{i=1}^{n} p_i.$$

上述性质可以为一些概率计算带来方便. 例如, 在第一章的生日问题中, 如果用A_{ij}表示第i 和第j个人同生日这一事件, 那么$\{A_{ij}, 1 \le i < j \le n\}$共有$\binom{n}{2}$个事件, 每个事件发生的概率为$\mathrm{P}(A_{ij}) = \frac{1}{365}$. 用Poisson分布$P(\lambda)$, $\lambda = \binom{n}{2} \times \frac{1}{365}$, 来近似这些事件发生的次数$\xi$的分布, 那么得$n$个人生日互不相同的概率为

上述性质可以为一些概率计算带来方便. 例如, 在第一章的生日问题中, 如果用A_{ij}表示第i 和第j个人同生日这一事件, 那么$\{A_{ij}, 1 \le i < j \le n\}$共有$\binom{n}{2}$个事件, 每个事件发生的概率为$\mathsf{P}(A_{ij}) = \frac{1}{365}$. 用Poisson分布$P(\lambda)$, $\lambda = \binom{n}{2} \times \frac{1}{365}$, 来近似这些事件发生的次数$\xi$的分布, 那么得$n$个人生日互不相同的概率为

$$\mathsf{P}(\xi = 0) \approx \frac{\lambda^0}{0!} e^{-\lambda} = \exp\left\{-\frac{n(n-1)}{2 \times 365}\right\}.$$

从而, 至少有2人生日相同的概率近似等于

$$1 - \exp\left\{-\frac{n(n-1)}{2 \times 365}\right\}.$$

5. 几何分布(geometric distribution)

5. 几何分布(geometric distribution)

定义

若随机变量ξ的分布列为

$$P(\xi = k) = q^{k-1}p,\ \ p, q > 0,\ p + q = 1,\ \ k = 1, 2, \cdots,$$

则称ξ服从几何分布.

5. 几何分布(geometric distribution)

定义

若随机变量ξ的分布列为

$$P(\xi = k) = q^{k-1}p, \ \ p, q > 0, \ p + q = 1, \ \ k = 1, 2, \cdots,$$

则称ξ服从几何分布.

在Bernoulli概型中，若一次试验成功的概率为p，则直到首次成功的试验次数服从几何分布. 即$r = 1$时的Pascal分布即为几何分布.

几何分布的性质：无记忆性

几何分布的性质: 无记忆性

若Bernoulli试验中前m次失败, 则从第$m+1$次开始直到首次成功的试验次数也服从几何分布(好像把前面的m次失败“忘记”了). 即,

$$\mathrm{P}(\xi = m+k|\xi > m) = \mathrm{P}(\xi = k) = q^{k-1}p.$$

几何分布的性质: 无记忆性

若Bernoulli试验中前m次失败, 则从第$m+1$次开始直到首次成功的试验次数也服从几何分布(好像把前面的m次失败“忘记”了). 即,

$$\mathsf{P}(\xi = m+k \mid \xi > m) = \mathsf{P}(\xi = k) = q^{k-1}p.$$

事实上,

$$\begin{aligned}
\mathsf{P}(\xi = m+k \mid \xi > m) =& \frac{\mathsf{P}(\xi = m+k, \xi > m)}{\mathsf{P}(\xi > m)} \\
=& \frac{\mathsf{P}(\xi = m+k)}{\mathsf{P}(\xi > m)} \\
=& \frac{q^{m+k-1}p}{\sum_{i=1}^{\infty} \mathsf{P}(\xi = m+i)} \\
=& \frac{q^{m+k-1}p}{\sum_{i=1}^{\infty} q^{m+i-1}p} = \frac{q^{m+k-1}p}{q^m} \\
=& q^{k-1}p = \mathsf{P}(\xi = k).
\end{aligned}$$

定理

若ξ是取正整数的随机变量, 且具有无记忆性, 则ξ服从几何分布.

*证: ξ具有无记忆性, 记

$$p = \mathsf{P}(\xi = k+1 | \xi > k), \; q_k = \mathsf{P}(\xi > k), \; p_k = \mathsf{P}(\xi = k).$$

那么$p_{k+1} = q_k - q_{k+1}$, 而且在已知$\xi > k$的条件下$\xi = k+1$的条件概率为p_{k+1}/q_k. 因此

$$\frac{p_{k+1}}{q_k} = p.$$

即

$$\frac{q_{k+1}}{q_k} = 1 - p.$$

注意到$q_0 = 1$, 那么$q_k = (1-p)^k$. 因此,

$$p_k = (1-p)^{k-1} p, \quad k = 1, 2, \cdots,$$

这正是几何分布的分布列.

6. 超几何分布(hype-geometric distribution)

6. 超几何分布(hype-geometric distribution)

定义

若随机变量ξ的分布列为

$$P(\xi = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \ n \le N, M \le N, \ k = 0, 1, \cdots, \min(n, M).$$

则称ξ服从超几何分布.

6. 超几何分布(hype-geometric distribution)

定义

若随机变量ξ的分布列为

$$P(\xi=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\ n\le N, M\le N,\ k=0,1,\cdots,\min(n,M).$$

则称ξ服从超几何分布.

超几何分布的一个典型例子: 在产品质量的不放回抽样中, 若N件产品中有M件次品, 则抽检n件时所得次品数服从超几何分布.

超几何分布的性质:

超几何分布的性质:

若n,k不变, $N\to\infty, M/N\to p$, 则超几何分布可用二项分布来逼近. 即,

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\to\binom{n}{k}p^kq^{n-k},\quad N\to\infty.$$

超几何分布的性质:

若n,k不变, $N\to\infty, M/N\to p$, 则超几何分布可用二项分布来逼近. 即,

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\to\binom{n}{k}p^kq^{n-k},\quad N\to\infty.$$

事实上,

$$\begin{aligned}&\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\\=&\frac{M!}{k!(M-k)!}\cdot\frac{(N-M)!}{(n-k)!(N-M-n+k)!}\cdot\frac{n!(N-n)!}{N!}\\=&\binom{n}{k}\cdot\frac{(M-k+1)\cdots M}{(N-k+1)\cdots N}\cdot\frac{(N-M-n+k+1)\cdots(N-M)}{(N-n+1)\cdots(N-k)}\\\to&\binom{n}{k}p^kq^{n-k},\quad N\to\infty.\end{aligned}$$

超几何分布的推广:

超几何分布的推广:

假设某N件产品中包含一、二、三级产品各为n_1, n_2, $N-n_1-n_2$件. 现从中抽查r件, 那么包含k_1件一级产品、k_2件二级产品、$r-k_1-k_2$件三级产品的概率为

$$\frac{\binom{n_1}{k_1}\binom{n_2}{k_2}\binom{N-n_1-n_2}{r-k_1-k_2}}{\binom{N}{r}}, \quad k_1 \le \min\{n_1, r\}, k_2 \le \min\{n_2, r\},$$
$$r-k_1-k_2 \le \min\{N-n_1-n_2, r\}.$$

例5 一套扑克52张, 由四种花色组成, 每种花色各13张. 那么一手13张扑克中有5张黑桃、4张红心、3张方块、1张梅花的概率为

例5 一套扑克52张, 由四种花色组成, 每种花色各13张. 那么一手13张扑克中有5张黑桃、4张红心、3张方块、1张梅花的概率为

$$\frac{\binom{13}{5}\binom{13}{4}\binom{13}{3}\binom{13}{1}}{\binom{52}{13}}.$$

分布函数与连续型随机变量

2.2.1 分布函数((cumulative) distribution function/CDF)

分布函数与连续型随机变量

2.2.1 **分布函数**((cumulative) distribution function/CDF)

1. 定义

分布函数与连续型随机变量

2.2.1 **分布函数**((cumulative) distribution function/CDF)

1. 定义

为何需要分布函数?

分布函数与连续型随机变量

2.2.1 **分布函数**((cumulative) distribution function/CDF)

1. 定义

为何需要分布函数? 用分布列来表示概率分布只适用于离散型随机变量. 若一随机变量的取值为一连续区间的一切值, 则无法一一罗列这些值及其概率, 此时分布列这一工具就失去了意义. 我们需要寻找新的能刻画随机变量的概率分布的工具, 使得它适用于一切随机变量.

因为对于随机变量来说, 我们关心的是随机变量落在Borel集上的概率. 而Borel集可由左开右闭区间通过(有限的或可列无穷的)并、交、逆运算得到. 自然地, 我们关心

$$\mathsf{P}(a < \xi \leq b)$$

这种类型的概率. 而

$$\mathsf{P}(a < \xi \leq b) = \mathsf{P}(\xi \leq b) - \mathsf{P}(\xi \leq a),$$

所以我们只需关心

$$\mathsf{P}(\xi \leq x)$$

这一类型的概率即可.

定义（分布函数）

设ξ为概率空间$(\Omega, \mathscr{F}, P)$上的随机变量. 我们称

$$F(x) = P(\xi \leq x), \quad -\infty < x < \infty$$

为随机变量ξ的分布函数.

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$\mathsf{P}(a < \xi \leq b) =$$

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$\mathsf{P}(a < \xi \leq b) = F(b) - F(a);$$

有了分布函数这一工具, 对于任意的Borel集B, 概率P$(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$\mathrm{P}(a < \xi \leq b) = F(b) - F(a);$$
$$\mathrm{P}(\xi < a) =$$

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$\mathsf{P}(a < \xi \leq b) = F(b) - F(a);$$
$$\mathsf{P}(\xi < a) = \lim_{b \to a-0} \mathsf{P}(\xi \leq b) = F(a-0);$$

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$\mathsf{P}(a < \xi \leq b) = F(b) - F(a);$$
$$\mathsf{P}(\xi < a) = \lim_{b \to a-0} \mathsf{P}(\xi \leq b) = F(a-0);$$
$$\mathsf{P}(\xi = a) =$$

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$\mathsf{P}(a < \xi \le b) = F(b) - F(a);$$
$$\mathsf{P}(\xi < a) = \lim_{b \to a-0} \mathsf{P}(\xi \le b) = F(a-0);$$
$$\mathsf{P}(\xi = a) = \mathsf{P}(\xi \le a) - \mathsf{P}(\xi < a) = F(a) - F(a-0);$$

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$\begin{aligned}
&\mathsf{P}(a < \xi \leq b) = F(b) - F(a);\\
&\mathsf{P}(\xi < a) = \lim_{b \to a-0} \mathsf{P}(\xi \leq b) = F(a-0);\\
&\mathsf{P}(\xi = a) = \mathsf{P}(\xi \leq a) - \mathsf{P}(\xi < a) = F(a) - F(a-0);\\
&\mathsf{P}(\xi > a) =
\end{aligned}$$

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$\begin{aligned}
&\mathsf{P}(a < \xi \leq b) = F(b) - F(a);\\
&\mathsf{P}(\xi < a) = \lim_{b \to a-0} \mathsf{P}(\xi \leq b) = F(a-0);\\
&\mathsf{P}(\xi = a) = \mathsf{P}(\xi \leq a) - \mathsf{P}(\xi < a) = F(a) - F(a-0);\\
&\mathsf{P}(\xi > a) = 1 - \mathsf{P}(\xi \leq a) = 1 - F(a);
\end{aligned}$$

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$
\begin{aligned}
&\mathsf{P}(a < \xi \leq b) = F(b) - F(a);\\
&\mathsf{P}(\xi < a) = \lim_{b \to a-0} \mathsf{P}(\xi \leq b) = F(a-0);\\
&\mathsf{P}(\xi = a) = \mathsf{P}(\xi \leq a) - \mathsf{P}(\xi < a) = F(a) - F(a-0);\\
&\mathsf{P}(\xi > a) = 1 - \mathsf{P}(\xi \leq a) = 1 - F(a);\\
&\mathsf{P}(a < \xi < b) =
\end{aligned}
$$

有了分布函数这一工具, 对于任意的Borel集B, 概率$\mathsf{P}(\xi \in B)$ 都可以用分布函数来表示了. 例如:

$$
\begin{aligned}
&\mathsf{P}(a < \xi \leq b) = F(b) - F(a);\\
&\mathsf{P}(\xi < a) = \lim_{b \to a-0} \mathsf{P}(\xi \leq b) = F(a-0);\\
&\mathsf{P}(\xi = a) = \mathsf{P}(\xi \leq a) - \mathsf{P}(\xi < a) = F(a) - F(a-0);\\
&\mathsf{P}(\xi > a) = 1 - \mathsf{P}(\xi \leq a) = 1 - F(a);\\
&\mathsf{P}(a < \xi < b) = \mathsf{P}(\xi < b) - \mathsf{P}(\xi \leq a) = F(b-0) - F(a).
\end{aligned}
$$

例1 设随机变量ξ服从$0-1(p)$分布. 请写出它的分布函数, 并计算P$(-1<\xi<0.5)$.

例1 设随机变量ξ服从$0-1(p)$分布. 请写出它的分布函数, 并计算P$(-1<\xi<0.5)$.

解: 我们需要对x的取值分类讨论.

例1 设随机变量ξ服从$0-1(p)$分布. 请写出它的分布函数, 并计算$\mathsf{P}(-1<\xi<0.5)$.

解: 我们需要对x的取值分类讨论. 当$x<0$时,

$$F(x)=\mathsf{P}(\xi\leq x)=\mathsf{P}(\phi)=0;$$

当$0\leq x<1$时,

$$F(x)=\mathsf{P}(\xi\leq x)=\mathsf{P}(\xi=0)=q;$$

当$x\geq 1$时,

$$F(x)=\mathsf{P}(\xi\leq x)=\mathsf{P}(\xi=0)+\mathsf{P}(\xi=1)=p+q=1.$$

因此, ξ的分布函数为

$$F(x) = \begin{cases} 0, & x < 0, \\ q, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

因此, ξ的分布函数为

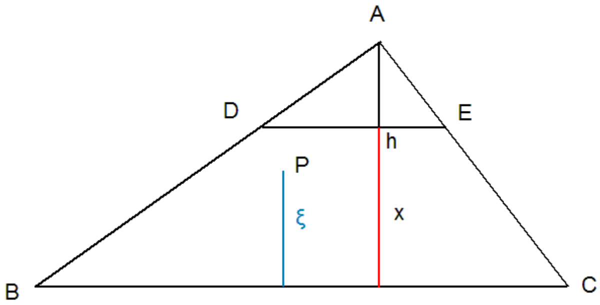
$$F(x) = \begin{cases} 0, & x < 0, \\ q, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

$$\mathsf{P}(-1 < \xi < 0.5) = F(0.5-0) - F(-1) = q - 0 = q.$$

例2 在三角形ΔABC内任取一点P, P到BC的距离为ξ. 求ξ的分布函数.

例2 在三角形ΔABC内任取一点P, P到BC的距离为ξ. 求ξ的分布函数.

解: 设BC边上的高为h, $DE \parallel BC$, 距离为x. 依图求解:



当$x<0$时，显然

$$F(x)=\mathsf{P}(\xi\le x)=0;$$

当$0\le x<h$时，$\{\xi\le x\}$与事件$\{P$落在梯形$DBCE$内$\}$是等价的.
因此由几何概率的定义得

$$F(x)=\mathsf{P}(\xi\le x)=\frac{\text{梯形}DBCE\text{的面积}}{\Delta ABC\text{的面积}}=1-(1-x/h)^2;$$

当$x\ge h$时，$\{\xi\le x\}$表示P在ΔABC内任意取. 因此

$$F(x)=\mathsf{P}(\xi\le x)=1.$$

当$x<0$时，显然
$$F(x)=\mathsf{P}(\xi\le x)=0;$$
当$0\le x<h$时，$\{\xi\le x\}$与事件$\{P$落在梯形$DBCE$内$\}$是等价的. 因此由几何概率的定义得
$$F(x)=\mathsf{P}(\xi\le x)=\frac{\text{梯形}DBCE\text{的面积}}{\Delta ABC\text{的面积}}=1-(1-x/h)^2;$$
当$x\ge h$时，$\{\xi\le x\}$表示P在ΔABC内任意取. 因此
$$F(x)=\mathsf{P}(\xi\le x)=1.$$
综上所述，ξ的分布函数为
$$F(x)=\begin{cases}0, & x<0,\\ 1-(1-x/h)^2, & 0\le x<h,\\ 1, & x\ge h.\end{cases}$$

2. 性质

2. 性质

分布函数具有下列性质:

(1) 若$a \leq b$, 则$F(a) \leq F(b)$; (单调不减性)
(2) $F(-\infty) := \lim\limits_{x \to -\infty} F(x) = 0,\ F(+\infty) := \lim\limits_{x \to +\infty} F(x) = 1;$
(3) $F(x+0) = \lim\limits_{\varepsilon \searrow 0} F(x+\varepsilon) = F(x)$. (右连续性)

2. 性质

分布函数具有下列性质:

(1) 若$a \leq b$, 则$F(a) \leq F(b)$; (单调不减性)
(2) $F(-\infty) := \lim\limits_{x \to -\infty} F(x) = 0,\ F(+\infty) := \lim\limits_{x \to +\infty} F(x) = 1$;
(3) $F(x+0) = \lim\limits_{\varepsilon \searrow 0} F(x+\varepsilon) = F(x)$. (右连续性)

证明: (1) 这是因为$F(b) - F(a) = \mathsf{P}(a < \xi \leq b) \geq 0$.

(2) 由于$F(x)$单调有界, 所以存在极限

$$F(-\infty) = \lim_{n\to\infty} F(-n).$$

(2) 由于$F(x)$单调有界, 所以存在极限

$$F(-\infty) = \lim_{n\to\infty} F(-n).$$

因为

$$\{\xi \le -(n+1)\} \subset \{\xi \le -n\}, \quad \bigcap_{n=1}^{\infty} \{\xi \le -n\} = \phi,$$

所以由概率的连续性定理得

$$\begin{aligned}\lim_{n\to\infty} F(-n) &= \lim_{n\to\infty} \mathsf{P}(\xi \le -n) \\ &= \mathsf{P}\Big(\bigcap_{n=1}^{\infty} \{\xi \le -n\}\Big) \\ &= \mathsf{P}(\phi) = 0.\end{aligned}$$

同理可得$F(+\infty) = 1$.

(3) 由$F(x)$的单调有界性, 我们只需证

$$\lim_{n\to\infty} F(x+1/n) = F(x).$$

(3) 由$F(x)$的单调有界性, 我们只需证

$$\lim_{n\to\infty} F(x+1/n) = F(x).$$

因为

$$\{\xi \le x+1/n\} \subset \{\xi \le x+1/(n-1)\}, \quad \bigcap_{n=1}^{\infty}\{\xi \le x+1/n\} = \{\xi \le x\},$$

所以

$$\begin{aligned}\lim_{n\to\infty} F(x+1/n) &= \lim_{n\to\infty} \mathsf{P}(\xi \le x+1/n)\\ &= \mathsf{P}\Big(\bigcap_{n=1}^{\infty}\{\xi \le x+1/n\}\Big)\\ &= \mathsf{P}(\xi \le x)\\ &= F(x).\end{aligned}$$

定理

函数$F(x)$为随机变量的分布函数的充要条件是它满足分布函数性质$(1)-(3)$.

证：必要性已证. 充分性见课本page85例2.30.

例3 设随机变量的分布函数如下，试确定常数a和b.

$$F(x)=\begin{cases}0, & x\le -1,\\ a+b\arcsin x, & -1<x\le 1,\\ 1, & x>1.\end{cases}$$

例3 设随机变量的分布函数如下，试确定常数a和b.

$$F(x)=\begin{cases} 0, & x\le -1, \\ a+b\arcsin x, & -1<x\le 1, \\ 1, & x>1. \end{cases}$$

解：利用分布函数的性质来确定常数a和b.

例3 设随机变量的分布函数如下，试确定常数a和b.

$$F(x)=\begin{cases}0, & x\le -1,\\ a+b\arcsin x, & -1<x\le 1,\\ 1, & x>1.\end{cases}$$

解: 利用分布函数的性质来确定常数a和b. 由右连续性,

$$\begin{cases}F(-1+0)=F(-1),\\ F(1+0)=F(1).\end{cases}$$

即,

$$\begin{cases}a-b\pi/2=0,\\ a+b\pi/2=1.\end{cases}$$

例3 设随机变量的分布函数如下，试确定常数a和b.

$$F(x)=\begin{cases}0, & x\le -1,\\ a+b\arcsin x, & -1<x\le 1,\\ 1, & x>1.\end{cases}$$

解: 利用分布函数的性质来确定常数a和b. 由右连续性,

$$\begin{cases}F(-1+0)=F(-1),\\ F(1+0)=F(1).\end{cases}$$

即,

$$\begin{cases}a-b\pi/2=0,\\ a+b\pi/2=1.\end{cases}$$

解得: $a=1/2, b=1/\pi$.

3. 离散型随机变量的分布函数

3. 离散型随机变量的分布函数

设ξ的分布列为

ξ	x_1	x_2	$\cdots$	x_k	$\cdots$
P	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_k)$	$\cdots$

且$x_1 < x_2 < \cdots < x_k < \cdots$. 则$\xi$的分布函数为

3. 离散型随机变量的分布函数

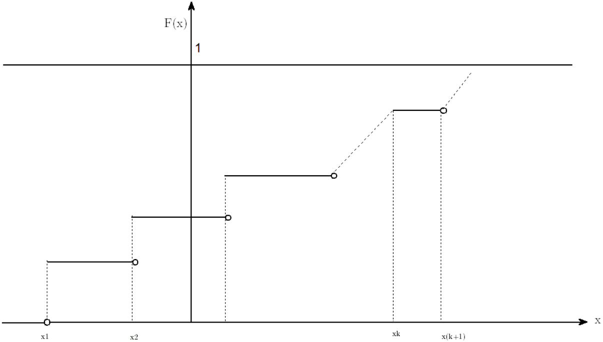
设ξ的分布列为

ξ	x_1	x_2	$\cdots$	x_k	$\cdots$
P	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_k)$	$\cdots$

且$x_1 < x_2 < \cdots < x_k < \cdots$. 则$\xi$的分布函数为

$$F(x) = \begin{cases} 0, & x < x_1, \\ p(x_1), & x_1 \le x < x_2, \\ \vdots & \vdots \\ \sum\limits_{i \le k} p(x_i), & x_k \le x < x_{k+1}, \\ \vdots & \vdots \end{cases}$$

图像如下(注意空心点的位置):



2.2.2 连续型随机变量(continuous random variables)及概率密度函数(probability density function/PDF)

2.2.2 连续型随机变量(continuous random variables)及概率密度函数(probability density function/PDF)

定义（连续型随机变量）

若随机变量ξ可取某个区间(有限或无限)中的一切值, 并且存在某个非负的可积函数$p(x)$, 使分布函数$F(x)$满足

$$F(x) = \int_{-\infty}^{x} p(y)\mathrm{d}y,$$

则称ξ为连续型随机变量, 称$p(x)$为ξ的(概率)密度函数.

连续型随机变量的分布函数是连续的.

连续型随机变量除了具有普通随机变量的分布函数的性质外, 还具有如下的性质:

连续型随机变量除了具有普通随机变量的分布函数的性质外，还具有如下的性质:

(1) $F(x)$是连续函数，在$p(x)$的连续点上$F(x)$可导，且

$$F^{'}(x)=p(x).$$

连续型随机变量除了具有普通随机变量的分布函数的性质外, 还具有如下的性质:

(1) $F(x)$是连续函数, 在$p(x)$的连续点上$F(x)$可导, 且

$$F^{'}(x) = p(x).$$

注: 因为$p(x)$是可积的, 所以它是几乎处处连续的. 因此, 除了一个测度为零的点集外, 连续型随机变量的分布函数和密度函数是一一对应的.

(2) ξ落在任何一个一维Borel集B上的概率可通过求定积分来确定:

$$\begin{aligned} \mathsf{P}(a < \xi \le b) =& F(b) - F(a) \\ =& \int_{-\infty}^{b} p(y)\mathrm{d}y - \int_{-\infty}^{a} p(y)\mathrm{d}y \\ =& \int_{a}^{b} p(y)\mathrm{d}y, \end{aligned}$$

因此,

$$\mathsf{P}(\xi \in B) = \int_B p(y)\mathrm{d}y.$$

(2) ξ落在任何一个一维Borel集B上的概率可通过求定积分来确定:

$$\begin{aligned}\mathsf{P}(a<\xi\le b) =& F(b)-F(a)\\ &=\int_{-\infty}^{b} p(y)\mathrm{d}y-\int_{-\infty}^{a} p(y)\mathrm{d}y\\ &=\int_{a}^{b} p(y)\mathrm{d}y,\end{aligned}$$

因此,

$$\mathsf{P}(\xi\in B)=$$

(2) ξ落在任何一个一维Borel集B上的概率可通过求定积分来确定:

$$\begin{aligned}\mathsf{P}(a<\xi\leq b)=&F(b)-F(a)\\=&\int_{-\infty}^{b}p(y)\mathrm{d}y-\int_{-\infty}^{a}p(y)\mathrm{d}y\\=&\int_{a}^{b}p(y)\mathrm{d}y,\end{aligned}$$

因此,

$$\mathsf{P}(\xi\in B)=\int_{B}p(y)\mathrm{d}y.$$

(3) 对任意常数c,

$$\mathsf{P}(\xi = c) =$$

(3) 对任意常数c,

$$\mathsf{P}(\xi = c) = F(c) - F(c-0) = \lim_{h \searrow 0} \int_{c-h}^{c} p(y)\mathsf{d}y = 0.$$

因此, 连续型随机变量等于任何一个常数的概率都为0. 这与离散型随机变量有本质的区别.

注: (1) 虽然$\mathsf{P}(\xi = c) = 0$, 但$\{\xi = c\}$是一个可能发生的事件. 所以

$$\mathsf{P}(A) = 0 \nRightarrow A = \phi.$$

注: (1) 虽然$\mathsf{P}(\xi = c) = 0$, 但$\{\xi = c\}$是一个可能发生的事件. 所以

$$\mathsf{P}(A) = 0 \nRightarrow A = \phi.$$

同理,

$$\mathsf{P}(A) = 1 \nRightarrow A = \Omega.$$

注: (1) 虽然P$(\xi = c) = 0$, 但$\{\xi = c\}$是一个可能发生的事件. 所以

$$\mathsf{P}(A) = 0 \nRightarrow A = \phi.$$

同理,

$$\mathsf{P}(A) = 1 \nRightarrow A = \Omega.$$

(2) 因为连续型随机变量等于任何一个实数的概率为0, 所以

$$\mathsf{P}(x < \xi \leq y) = \mathsf{P}(x \leq \xi \leq y) = \mathsf{P}(x \leq \xi < y) = \mathsf{P}(x < \xi < y).$$

密度函数的性质:

密度函数的性质:

- 非负性: $p(y) \geq 0$;
- 规范性: $\int_{-\infty}^{\infty} p(y)\mathrm{d}y = 1$.

密度函数的性质:

- 非负性: $p(y) \geq 0$;
- 规范性: $\int_{-\infty}^{\infty} p(y)\mathrm{d}y = 1$.

注: (1) 由于在$p(x)$的连续点x处,

$$p(x)\Delta x \approx \int_{x}^{x+\Delta x} p(y)\mathrm{d}y = \mathsf{P}(x < \xi \leq x + \Delta x).$$

即

$$p(x) \approx \frac{\mathsf{P}(x < \xi \leq x + \Delta x)}{\Delta x}.$$

因此概率密度$p(x)$的数值反映了随机变量取x的邻近的值的概率大小.

随机变量的分类: 离散型、连续型、其他.

随机变量的分类: 离散型、连续型、其他.

例:

$$F(x)=\begin{cases}0, & x<0,\\ (1+x)/2, & 0\le x<1,\\ 1, & x\ge 1\end{cases}$$

是分布函数(因为满足分布函数的三个性质), 它对应的随机变量既不是离散型的, 也不是连续型的(因为$F(x)$不连续).

今后, 对于一般的随机变量, 我们将采用分布函数去描述随机变量的特性.

对于离散型随机变量, 分布函数与分布列可以互相确定, 我们通常用分布列去描述随机变量的特性.

对于连续型随机变量, 分布函数与密度函数可以互相确定, 我们通常用密度函数去描述随机变量的特性.

2.2.3 常见的连续型随机变量

2.2.3 常见的连续型随机变量

1. 均匀分布(uniform distribution)

2.2.3 常见的连续型随机变量

1. 均匀分布(uniform distribution)

定义 (均匀分布)

若随机变量ξ具有密度函数

$$p(x)=\begin{cases}\frac{1}{b-a}, & a<x<b,\\ 0, & \text{其他}.\end{cases} \tag{2.2.1}$$

则称ξ服从(a,b)上的均匀分布. 记为$\xi\sim U(a,b)$.

2.2.3 常见的连续型随机变量

1. 均匀分布(uniform distribution)

定义 (均匀分布)

若随机变量ξ具有密度函数

$$p(x) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & \text{其他}. \end{cases} \tag{2.2.1}$$

则称ξ服从(a,b)上的均匀分布. 记为$\xi \sim U(a,b)$.

容易验证(2.2.1)式满足密度函数的两个条件.

我们来确定$\xi \sim U(a,b)$的分布函数.

我们来确定$\xi \sim U(a,b)$的分布函数.

当$x \leq a$时, 显然$F(x) = \mathsf{P}(\xi \leq x) = 0$;

我们来确定$\xi \sim U(a,b)$的分布函数.

当$x \le a$时, 显然$F(x) = \mathsf{P}(\xi \le x) = 0$;

当$a < x < b$时,

$$F(x) = \int_{-\infty}^{x} p(y)\mathrm{d}y = \int_{a}^{x} \frac{1}{b-a}\mathrm{d}y = \frac{x-a}{b-a};$$

我们来确定$\xi \sim U(a,b)$的分布函数.

当$x \le a$时，显然$F(x) = \mathsf{P}(\xi \le x) = 0$;

当$a < x < b$时，

$$F(x) = \int_{-\infty}^{x} p(y)\mathrm{d}y = \int_{a}^{x} \frac{1}{b-a}\mathrm{d}y = \frac{x-a}{b-a};$$

当$x \ge b$时，

$$F(x) = \int_{-\infty}^{x} p(y)\mathrm{d}y = \int_{a}^{b} \frac{1}{b-a}\mathrm{d}y = 1.$$

我们来确定$\xi \sim U(a,b)$的分布函数.

当$x \le a$时, 显然$F(x) = \mathsf{P}(\xi \le x) = 0$;

当$a < x < b$时,

$$F(x) = \int_{-\infty}^{x} p(y)\mathrm{d}y = \int_{a}^{x} \frac{1}{b-a}\mathrm{d}y = \frac{x-a}{b-a};$$

当$x \ge b$时,

$$F(x) = \int_{-\infty}^{x} p(y)\mathrm{d}y = \int_{a}^{b} \frac{1}{b-a}\mathrm{d}y = 1.$$

综上所述, ξ的分布函数为

$$F(x) = \begin{cases} 0, & x \le a, \\ \frac{x-a}{b-a}, & a < x < b, \\ 1, & x \ge b. \end{cases}$$

均匀分布的性质:

均匀分布的性质:

若$\xi \sim U(a,b)$, 区间$(c, c+l) \subset (a,b)$, 则

$$\mathsf{P}(c < \xi < c+l) =$$

均匀分布的性质:

若$\xi \sim U(a,b)$, 区间$(c, c+l) \subset (a,b)$, 则

$$\mathsf{P}(c<\xi<c+l)=\int_c^{c+l}\frac{1}{b-a}\mathrm{d}y=\frac{l}{b-a},\quad 与c无关.$$

所以, ξ在(a,b)中取值落在某一区间内的概率与这个区间的测度成正比, 与起点和终点无关. 粗略地讲, 就是(a,b)中所有点的重要性是“均匀”的.

2. 正态分布(normal distribution)

2. 正态分布(normal distribution)

定义 (正态分布)

若随机变量ξ的概率密度函数为

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}},\quad -\infty<x<\infty, \tag{2.2.2}$$

则称ξ服从参数为(a,σ^2)的正态分布$(-\infty<a<\infty,\sigma>0)$. 记作$\xi\sim N(a,\sigma^2)$.

正态分布又被称为误差分布或高斯(Gauss)分布.

我们来说明(2.2.2)定义的$p(x)$的确是密度函数.

我们来说明(2.2.2)定义的$p(x)$的确是密度函数.

$p(x) \geq 0$是显然的. 为了说明$\int_{-\infty}^{\infty} p(x)\mathrm{d}x = 1$, 记$I = \int_{-\infty}^{\infty} p(x)\mathrm{d}x > 0$. 则

$$
\begin{aligned}
I^2 =& \left(\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}}\mathrm{d}x\right)^2 \\
=& \left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\mathrm{d}t\right)^2 \quad (\text{令}t = \frac{x-a}{\sigma}) \\
=& \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{t^2+s^2}{2}}\mathrm{d}t\mathrm{d}s \\
=& \frac{1}{2\pi}\int_{0}^{2\pi}\int_{0}^{\infty} e^{-\frac{r^2}{2}} r\mathrm{d}r\mathrm{d}\theta \ (\text{极坐标变换}) \\
=& \int_{0}^{\infty} e^{-\frac{r^2}{2}} r\mathrm{d}r = 1.
\end{aligned}
$$

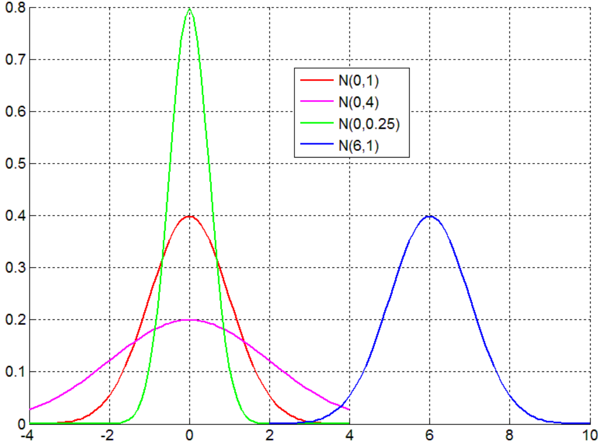
所以$I = 1$.

正态分布$N(a,\sigma^2)$的分布函数:

正态分布$N(a,\sigma^2)$的分布函数:

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-a)^2}{2\sigma^2}}\mathrm{d}t,\quad -\infty<x<\infty.$$

正态分布的密度函数的图像:



正态分布$N(a,\sigma^2)$的性质:

正态分布$N(a,\sigma^2)$的性质:

- $p(x)$关于直线$x=a$对称;
- 当$x>a$时, $p(x)$单调递减; 当$x<a$时, $p(x)$单调递增; $x\to\pm\infty$ 时, $p(x)\to 0$; $x=a$ 时, $p(x)$有最大值$\frac{1}{\sqrt{2\pi}\sigma}$; $a\pm\sigma$为$p(x)$的两个拐点. σ越大, $p(x)$的图像越扁平, ξ取值离开a点远的概率也越大; σ越小, $p(x)$的图像越陡峭, ξ取值越集中在点a附近. (称a为位置参数, σ为尺度参数)

正态分布是概率论中最重要的分布:

正态分布是概率论中最重要的分布:

(1)正态分布是自然界中最常见的一种分布, 如测量的误差; 炮弹弹落点的分布; 人的身高、体重等; 农作物的收获量; 工厂产品的尺寸: 直径、宽度、高度、长度$\cdots$都近似服从正态分布;

正态分布是概率论中最重要的分布:

(1)正态分布是自然界中最常见的一种分布, 如测量的误差; 炮弹弹落点的分布; 人的身高、体重等; 农作物的收获量; 工厂产品的尺寸: 直径、宽度、高度、长度$\cdots$都近似服从正态分布;

(2)一般说来, 若影响某一数量指标的随机因素很多, 而每个因素所起的作用不太大, 则这个指标服从正态分布, 这点可以利用概率论的极限定理来加以证明(见本书第四章);

正态分布是概率论中最重要的分布:

(1)正态分布是自然界中最常见的一种分布, 如测量的误差; 炮弹弹落点的分布; 人的身高、体重等; 农作物的收获量; 工厂产品的尺寸: 直径、宽度、高度、长度···都近似服从正态分布;

(2)一般说来, 若影响某一数量指标的随机因素很多, 而每个因素所起的作用不太大, 则这个指标服从正态分布, 这点可以利用概率论的极限定理来加以证明(见本书第四章);

(3)数理统计中一些重要的分布可以通过正态分布来导出.

标准正态分布:

标准正态分布:

称$a=0,\sigma=1$时的正态分布为标准正态分布, 记为$N(0,1)$. 它的密度函数的图像关于纵轴对称. 其密度函数和分布函数被特别记为$\varphi(x)$和$\Phi(x)$: 即,

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

$$\Phi(x)=\int_{-\infty}^{x}\varphi(t)\mathrm{d}t,\quad -\infty<x<\infty.$$

正态分布的计算:

正态分布的计算:

(1) $\xi \sim N(0,1)$时,

- 当$x \geq 0$时, 查正态分布表(附录C): 每隔一定数值可以查到对应的分布函数$\Phi(x)$ 的值(特别地, $\Phi(0)=0.5$); 在这些数值之间, 可以用线性插值法求得相应的函数值.

正态分布的计算:

(1) $\xi \sim N(0,1)$时,

- 当$x \geq 0$时, 查正态分布表(附录C): 每隔一定数值可以查到对应的分布函数$\Phi(x)$ 的值(特别地, $\Phi(0) = 0.5$); 在这些数值之间, 可以用线性插值法求得相应的函数值.
- 当$x < 0$时, 注意到

$$\Phi(x) + \Phi(-x) = 1,$$

因此可以先通过查表或线性插值得到$\Phi(-x)$, 再求得

$$\Phi(x) = 1 - \Phi(-x).$$

(2) $\xi \sim N(a, \sigma^2)$时,

(2) $\xi \sim N(a, \sigma^2)$时，记$\eta = \frac{\xi - a}{\sigma}$. η的分布函数为

(2) $\xi \sim N(a,\sigma^2)$时，记$\eta = \frac{\xi-a}{\sigma}$. η的分布函数为

$$
\begin{aligned}
\mathsf{P}(\eta \le x) =& \mathsf{P}(\xi \le a+\sigma x) \\
=& \int_{-\infty}^{a+\sigma x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-a)^2}{2\sigma^2}} \mathrm{d}t \\
=& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} \mathrm{d}u = \Phi(x).
\end{aligned}
$$

所以$\eta \sim N(0,1)$.

(2) $\xi \sim N(a, \sigma^2)$时，记$\eta = \frac{\xi - a}{\sigma}$. η的分布函数为

$$\begin{aligned}\mathrm{P}(\eta \le x) =& \mathrm{P}(\xi \le a + \sigma x) \\ =& \int_{-\infty}^{a+\sigma x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-a)^2}{2\sigma^2}} \mathrm{d}t \\ =& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} \mathrm{d}u = \Phi(x).\end{aligned}$$

所以$\eta \sim N(0, 1)$. 因此对于任意的$x < y$, 有

$$\begin{aligned}\mathrm{P}(x < \xi \le y) =& \mathrm{P}(\frac{x-a}{\sigma} < \eta \le \frac{y-a}{\sigma}) \\ =& \Phi(\frac{y-a}{\sigma}) - \Phi(\frac{x-a}{\sigma}).\end{aligned}$$

例4 设$\xi \sim N(0,1)$.

(1) 计算P$(-1 < \xi < 3)$;

(2) 已知P$(\xi < \lambda) = 0.9755$, 求λ.

例4 设$\xi \sim N(0,1)$.
(1) 计算$\mathrm{P}(-1 < \xi < 3)$;
(2) 已知$\mathrm{P}(\xi < \lambda) = 0.9755$, 求$\lambda$.

解: (1) 查正态分布表得

$$\begin{aligned}\mathrm{P}(-1 < x < 3) =& \Phi(3) - \Phi(-1) = \Phi(3) + \Phi(1) - 1 \\ =& 0.9987 + 0.8413 - 1 = 0.8400.\end{aligned}$$

例4 设$\xi \sim N(0,1)$.
(1) 计算P$(-1<\xi<3)$;
(2) 已知P$(\xi<\lambda)=0.9755$, 求λ.

解: (1) 查正态分布表得

$$\begin{aligned}\mathrm{P}(-1<x<3)=&\Phi(3)-\Phi(-1)=\Phi(3)+\Phi(1)-1\\=&0.9987+0.8413-1=0.8400.\end{aligned}$$

(2) $\Phi(\lambda)=0.9755$, 在$\Phi(1.96)=0.9750$与$\Phi(1.98)=0.9762$之间.
所以用线性插值得

$$\begin{aligned}\lambda\approx&1.96+\frac{\Phi(\lambda)-\Phi(1.96)}{\Phi(1.98)-\Phi(1.96)}\times(1.98-1.96)\\\approx&1.968.\end{aligned}$$

例5(“3σ”原则) 设$\xi \sim N(a,\sigma^2)$. 求

$$\mathrm{P}(|\xi-a|<\sigma),\ \mathrm{P}(|\xi-a|<2\sigma),\ \text{以及}\mathrm{P}(|\xi-a|<3\sigma).$$

例5("3σ" 原则) 设$\xi \sim N(a, \sigma^2)$. 求

$$\mathrm{P}(|\xi - a| < \sigma),\ \mathrm{P}(|\xi - a| < 2\sigma),\ \text{以及}\mathrm{P}(|\xi - a| < 3\sigma).$$

解: 注意到$\eta = (\xi - a)/\sigma \sim N(0, 1)$. 因此

$$\begin{aligned}
\mathrm{P}(|\xi - a| < \sigma) &= \mathrm{P}(|\eta| < 1) = 2\Phi(1) - 1 \approx 0.6827,\\
\mathrm{P}(|\xi - a| < 2\sigma) &= \mathrm{P}(|\eta| < 2) = 2\Phi(2) - 1 \approx 0.9545,\\
\mathrm{P}(|\xi - a| < 3\sigma) &= \mathrm{P}(|\eta| < 3) = 2\Phi(3) - 1 \approx 0.9973.
\end{aligned}$$

例5("3σ" 原则) 设$\xi \sim N(a,\sigma^2)$. 求

$$\mathsf{P}(|\xi-a|<\sigma),\ \mathsf{P}(|\xi-a|<2\sigma),\ \text{以及}\mathsf{P}(|\xi-a|<3\sigma).$$

解: 注意到$\eta=(\xi-a)/\sigma \sim N(0,1)$. 因此

$$\begin{aligned}
\mathsf{P}(|\xi-a|<\sigma) &= \mathsf{P}(|\eta|<1) = 2\Phi(1)-1 \approx 0.6827,\\
\mathsf{P}(|\xi-a|<2\sigma) &= \mathsf{P}(|\eta|<2) = 2\Phi(2)-1 \approx 0.9545,\\
\mathsf{P}(|\xi-a|<3\sigma) &= \mathsf{P}(|\eta|<3) = 2\Phi(3)-1 \approx 0.9973.
\end{aligned}$$

这说明ξ的值有99.73%的可能性落在$(a-3\sigma,a+3\sigma)$之间, 落在该区间之外的可能性几乎为零.

例6 从南郊某地乘车到北区火车站有两条路可走, 第一条路较短, 但交通拥挤, 所需时间$\tau \sim N(50,100)$; 第二条路线略长, 但意外阻塞较少, 所需时间$\xi \sim N(60,16)$.

(1) 若有70分钟可用, 问应走哪一条路?

(2) 若有65分钟可用, 问应走哪一条路?

例6 从南郊某地乘车到北区火车站有两条路可走, 第一条路较短, 但交通拥挤, 所需时间$\tau \sim N(50, 100)$; 第二条路线略长, 但意外阻塞较少, 所需时间$\xi \sim N(60, 16)$.
(1) 若有70分钟可用, 问应走哪一条路?
(2) 若有65分钟可用, 问应走哪一条路?

解: 应该走在允许时间内有较大概率赶到火车站的路线.

例6 从南郊某地乘车到北区火车站有两条路可走, 第一条路较短, 但交通拥挤, 所需时间$\tau \sim N(50, 100)$; 第二条路线略长, 但意外阻塞较少, 所需时间$\xi \sim N(60, 16)$.
(1) 若有70分钟可用, 问应走哪一条路?
(2) 若有65分钟可用, 问应走哪一条路?

解: 应该走在允许时间内有较大概率赶到火车站的路线.

(1) 走第一条路线能及时赶到火车站的概率为

$$\mathrm{P}(\tau \leq 70) = \Phi(\frac{70-50}{10}) = \Phi(2) = 0.9772;$$

走第二条路线能及时赶到火车站的概率为

$$\mathrm{P}(\xi \leq 70) = \Phi(\frac{70-60}{4}) = \Phi(2.5) = 0.9938.$$

因此在这种场合下, 应走第二条路线.

(2) 走第一条路线能及时赶到火车站的概率为

$$\mathsf{P}(\tau \le 65) = \Phi(\frac{65-50}{10}) = \Phi(1.5) = 0.9332;$$

走第二条路线能及时赶到火车站的概率为

$$\mathsf{P}(\xi \le 65) = \Phi(\frac{65-60}{4}) = \Phi(1.25) = 0.8944.$$

因此在这种场合下, 应走第一条路线.

例7 设$\xi \sim N(0,1)$. 证明

$$\mathsf{P}(|\xi| \geq x) \sim \sqrt{\frac{2}{\pi}}\frac{1}{x}e^{-\frac{x^2}{2}}, \quad x \to \infty.$$

其中$a(x) \sim b(x)(x \to \infty)$表示$\lim\limits_{x\to\infty} \dfrac{a(x)}{b(x)} = 1.$

例7 设$\xi \sim N(0,1)$. 证明

$$\mathsf{P}(|\xi| \geq x) \sim \sqrt{\frac{2}{\pi}} \frac{1}{x} e^{-\frac{x^2}{2}}, \quad x \to \infty.$$

其中$a(x) \sim b(x)(x \to \infty)$表示$\lim\limits_{x \to \infty} \dfrac{a(x)}{b(x)} = 1.$

证明: 由于对任意$x > 0$都有

$$\mathsf{P}(|\xi| \geq x) = \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-\frac{u^2}{2}} \mathrm{d}u,$$

所以只需证明

$$\int_x^\infty e^{-\frac{u^2}{2}} \mathrm{d}u \sim \frac{1}{x} e^{-\frac{x^2}{2}}, \quad x \to \infty.$$

易知

$$\int_x^\infty e^{-\frac{u^2}{2}} \mathrm{d}u = \int_x^\infty \frac{1}{u} \cdot u e^{-\frac{u^2}{2}} \mathrm{d}u = -\int_x^\infty \frac{1}{u} \mathrm{d}e^{-\frac{u^2}{2}},$$

利用分部积分可得,

$$\begin{aligned}\int_x^{\infty} e^{-\frac{u^2}{2}}\mathrm{d}u &= -\int_x^{\infty}\frac{1}{u}\mathrm{d}e^{-\frac{u^2}{2}}\\ &=\frac{1}{x}e^{-\frac{x^2}{2}}-\int_x^{\infty}\frac{1}{u^2}e^{-\frac{u^2}{2}}\mathrm{d}u\\ &\leq\frac{1}{x}e^{-\frac{x^2}{2}}.\end{aligned}$$

如果将上面的倒数第二步继续算下去，则又可得

$$
\begin{aligned}
\int_x^{\infty} e^{-\frac{u^2}{2}}\mathrm{d}u &= \frac{1}{x}e^{-\frac{x^2}{2}} - \int_x^{\infty} \frac{1}{u^2}e^{-\frac{u^2}{2}}\mathrm{d}u \\
&= \frac{1}{x}e^{-\frac{x^2}{2}} - \int_x^{\infty} \frac{1}{u^3}e^{-\frac{u^2}{2}}\mathrm{d}\frac{u^2}{2} \\
&= \frac{1}{x}e^{-\frac{x^2}{2}} + \int_x^{\infty} \frac{1}{u^3}\mathrm{d}e^{-\frac{u^2}{2}} \\
&= \left(\frac{1}{x} - \frac{1}{x^3}\right)e^{-\frac{x^2}{2}} + 3\int_x^{\infty} \frac{1}{u^4}e^{-\frac{u^2}{2}}\mathrm{d}u \\
&\geq \left(\frac{1}{x} - \frac{1}{x^3}\right)e^{-\frac{x^2}{2}}.
\end{aligned}
$$

综上所述, 得

$$\left(\frac{1}{x}-\frac{1}{x^3}\right)e^{-\frac{x^2}{2}} \le \int_x^{\infty} e^{-\frac{u^2}{2}}\mathrm{d}u \le \frac{1}{x}e^{-\frac{x^2}{2}}.$$

由此即可得到所要证明的结论.

3. 指数分布(exponential distribution)

3. 指数分布(exponential distribution)

定义（指数分布）

若随机变量ξ具有密度函数

$$p(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0, \end{cases} \tag{2.2.3}$$

其中$\lambda > 0$. 则称ξ服从参数为λ的指数分布. 记为$\xi \sim \text{Exp}(\lambda)$.

3. 指数分布(exponential distribution)

定义（指数分布）

若随机变量ξ具有密度函数

$$p(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0, \end{cases} \tag{2.2.3}$$

其中$\lambda > 0$. 则称ξ服从参数为λ的指数分布. 记为$\xi \sim \text{Exp}(\lambda)$.

容易验证(2.2.3)式满足密度函数的两个条件.

指数分布$\xi \sim \mathrm{Exp}(\lambda)$的分布函数:

指数分布$\xi \sim \mathrm{Exp}(\lambda)$的分布函数:

当$x < 0$时,

$$\mathrm{P}(\xi \le x) = \int_{-\infty}^{x} 0 \mathrm{d}x = 0;$$

当$x \ge 0$时,

$$\mathrm{P}(\xi \le x) = \int_{0}^{x} \lambda e^{-\lambda t} \mathrm{d}t = 1 - e^{-\lambda x}.$$

所以分布函数为

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

指数分布的重要性还表现在它具有类似几何分布的“无记忆性”：设随机变量$\xi \sim \mathrm{Exp}(\lambda)$, 则对任意的$s > 0, t > 0$,

$$\mathrm{P}(\xi \geq s + t | \xi \geq s) =$$

指数分布的重要性还表现在它具有类似几何分布的“无记忆性”：设随机变量$\xi \sim \mathsf{Exp}(\lambda)$, 则对任意的$s > 0, t > 0$,

$$\mathsf{P}(\xi \geq s + t | \xi \geq s) = \frac{\mathsf{P}(\xi \geq s + t)}{\mathsf{P}(\xi \geq s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t}.$$

指数分布的重要性还表现在它具有类似几何分布的"无记忆性"：设随机变量$\xi \sim \mathsf{Exp}(\lambda)$, 则对任意的$s > 0, t > 0$,

$$\mathsf{P}(\xi \geq s+t|\xi \geq s) = \frac{\mathsf{P}(\xi \geq s+t)}{\mathsf{P}(\xi \geq s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t}.$$

因此,

$$\mathsf{P}(\xi \geq s+t|\xi \geq s) = \mathsf{P}(\xi \geq t). \tag{2.2.4}$$

指数分布的重要性还表现在它具有类似几何分布的“无记忆性”：设随机变量$\xi \sim \mathrm{Exp}(\lambda)$, 则对任意的$s > 0, t > 0$,

$$\mathsf{P}(\xi \geq s + t | \xi \geq s) = \frac{\mathsf{P}(\xi \geq s + t)}{\mathsf{P}(\xi \geq s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t}.$$

因此,

$$\mathsf{P}(\xi \geq s + t | \xi \geq s) = \mathsf{P}(\xi \geq t). \tag{2.2.4}$$

假如把ξ解释为寿命, 则(2.2.4)表明, 如果已知某人/物/元件的年龄为s, 再活t年的概率与年龄s无关. 所以有时又风趣地称指数分布是“永远年轻”的.

指数分布有广泛的应用, 常用它来作为各种“寿命”分布的近似, 例如电子元件的寿命, 某些动物(例如乌龟)的寿命, 电话问题中的通话时间, 随机服务系统中的服务时间等都常假定服从指数分布.

定理

指数分布是具有性质(2.2.4)的唯一的连续型分布.

定理

指数分布是具有性质(*2.2.4*)的唯一的连续型分布.

*为了证明这个结论, 我们需要一个引理(证明略):

定理

指数分布是具有性质(2.2.4)的唯一的连续型分布.

*为了证明这个结论, 我们需要一个引理(证明略):

引理 (柯西)

若$f(x)$是连续函数(或单调函数), 且对一切x, y(或一切$x \geq 0$, $y \geq 0$)成立

$$f(x)f(y) = f(x+y),$$

则

$$f(x) = a^x,$$

其中$a \geq 0$为一常数.

*定理的证明: 记$G(x)=\mathsf{P}(\xi \geq x)$. 则由(2.2.4)得

$$G(s+t)=G(s)G(t)$$

对一切的$s \geq 0, t \geq 0$成立. 因为$G(x)$关于x单调, 所以由前面的引理可知

$$G(x)=a^x, \quad x \geq 0.$$

由于$G(x)$是概率, 故$0<a<1$, 可以写成$a=e^{-\lambda}$, 其中$\lambda>0$. 因此

$$F(x)=1-G(x)=1-e^{-\lambda x}, \quad x \geq 0.$$

再由分布函数的单调性可知当$x<0$时

$$F(x) \leq F(0)=0.$$

所以ξ服从指数分布.

4. Γ分布

4. Γ分布

定义 (Γ分布)

若随机变量ξ具有密度函数

$$p(x)=\begin{cases}0, & x<0,\\ \frac{\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x}, & x\geq 0,\end{cases} \tag{2.2.5}$$

其中$\lambda>0, r>0$, $\Gamma(r)=\int_0^{\infty}x^{r-1}e^{-x}\mathrm{d}x$. 则称$\xi$服从参数为$(\lambda,r)$的$\Gamma$分布. 记为$\xi\sim\Gamma(\lambda,r)$.

4. Γ分布

定义（Γ分布）

若随机变量ξ具有密度函数

$$p(x) = \begin{cases} 0, & x < 0, \\ \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, & x \geq 0, \end{cases} \tag{2.2.5}$$

其中$\lambda > 0, r > 0$, $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} \mathrm{d}x$. 则称$\xi$服从参数为$(\lambda, r)$的$\Gamma$分布. 记为$\xi \sim \Gamma(\lambda, r)$.

注: (1) 英文教材里通常记为$\xi \sim \Gamma(r, \lambda)$, 称$r$为形状参数(shape parameter), 称λ为逆尺度参数(inverse scale parameter)或称$1/\lambda$ 为尺度参数(scale parameter); (2) r为整数时的Γ分布也被称为埃尔朗(Erlang)分布, 它是丹麦科学家埃尔朗在研究电话问题时引进的, 他的这些研究开创了排队论这一学科; $r = 1$时的Γ分布即为指数分布Exp(λ); (3) 容易验证(2.2.5)式满足密度函数的两个条件.

5. *威布尔分布(Weibull distribution)

5. *威布尔分布(Weibull distribution)

若随机变量ξ的密度函数为

$$p(x)=\begin{cases}\frac{\alpha}{\sigma}\left(\frac{x-\mu}{\sigma}\right)^{\alpha-1}\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{\alpha}\right\}, & x>\mu,\\ 0, & x\le\mu,\end{cases}$$

则称ξ服从参数μ,σ和α的威布尔分布, 记作$\xi\sim \mathrm{Weib}(\mu,\sigma,\alpha)$.

5. *威布尔分布(Weibull distribution)

若随机变量ξ的密度函数为

$$p(x)=\begin{cases}\frac{\alpha}{\sigma}\left(\frac{x-\mu}{\sigma}\right)^{\alpha-1}\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{\alpha}\right\}, & x>\mu,\\ 0, & x\le\mu,\end{cases}$$

则称ξ服从参数μ,σ和α的威布尔分布，记作$\xi\sim \text{Weib}(\mu,\sigma,\alpha)$. 其分布函数为

$$F(x)=\begin{cases}1-\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{\alpha}\right\}, & x>\mu,\\ 0, & x\le\mu.\end{cases}$$

5. *威布尔分布(Weibull distribution)

若随机变量ξ的密度函数为

$$p(x)=\begin{cases}\frac{\alpha}{\sigma}\left(\frac{x-\mu}{\sigma}\right)^{\alpha-1}\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{\alpha}\right\}, & x>\mu,\\ 0, & x\le\mu,\end{cases}$$

则称ξ服从参数μ,σ和α的威布尔分布，记作$\xi\sim \text{Weib}(\mu,\sigma,\alpha)$. 其分布函数为

$$F(x)=\begin{cases}1-\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{\alpha}\right\}, & x>\mu,\\ 0, & x\le\mu.\end{cases}$$

当$\alpha=1,\mu=0$时，威布尔分布就是参数为$1/\sigma$指数分布. 威布尔分布在工程领域有广泛的应用，许多产品的使用寿命服从威布尔分布. 如今，威布尔分布广泛应用于与寿命有关的领域，如生存分析、保险精算、可靠性理论等.

6. *帕累托分布(Pareto distribution)

6. *帕累托分布(Pareto distribution)

若随机变量ξ的密度函数为

$$p(x)=\begin{cases}(\alpha-1)x_0^{\alpha-1}x^{-\alpha}, & x>x_0;\\ 0, & x\le x_0,\end{cases}$$

则称ξ服从帕累托分布, 其中参数$x_0>0,\alpha>1$.

6. *帕累托分布(Pareto distribution)

若随机变量ξ的密度函数为

$$p(x)=\begin{cases}(\alpha-1)x_0^{\alpha-1}x^{-\alpha}, & x>x_0;\\ 0, & x\le x_0,\end{cases}$$

则称ξ服从帕累托分布, 其中参数$x_0>0,\alpha>1$.

意大利经济学家帕累托首先引入这一分布来描述一个国家中家庭年收入的分布.

7. *β分布(Beta distribution)

7. *β分布(Beta distribution)

若随机变量ξ的密度函数为

$$p(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, & 0 \le x \le 1, \\ 0, & \text{其他}, \end{cases}$$

则称ξ服从参数为a和b的β分布, 记作$\xi \sim \beta(a,b)$, 其中$a, b > 0$, $B(a,b)$ 为β积分:

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x, \quad B(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}.$$

7. *β分布(Beta distribution)

若随机变量ξ的密度函数为

$$p(x)=\begin{cases}\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}, & 0\le x\le 1,\\ 0, & \text{其他},\end{cases}$$

则称ξ服从参数为a和b的β分布, 记作$\xi\sim\beta(a,b)$, 其中$a,b>0$, $B(a,b)$ 为β积分:

$$B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x,\ \ B(a,b)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}.$$

β分布可以用来为取值在有限区间上的随机现象建模. 当$a=b=1$时, $\beta(1,1)$分布就是区间$[0,1]$上的均匀分布. β分布与二项分布、Γ分布有密切联系.

8. *柯西分布(Cauchy distribution)

8. *柯西分布(Cauchy distribution)

定义（柯西分布）

若随机变量ξ具有密度函数

$$p(x) = \frac{1}{\pi}\frac{\lambda}{\lambda^2 + (x-\mu)^2}, \quad -\infty < x < \infty, \tag{2.2.6}$$

其中$\lambda > 0$, μ为任意常数. 则称ξ服从参数为(λ, μ)的柯西分布.

8. *柯西分布(Cauchy distribution)

定义 (柯西分布)

若随机变量ξ具有密度函数

$$p(x)=\frac{1}{\pi}\frac{\lambda}{\lambda^2+(x-\mu)^2},\quad -\infty<x<\infty, \tag{2.2.6}$$

其中$\lambda>0$, μ为任意常数. 则称ξ服从参数为(λ,μ)的柯西分布.

$\lambda=1,\mu=0$时的柯西分布的密度函数为

$$p(x)=\frac{1}{\pi}\frac{1}{1+x^2},\quad -\infty<x<\infty.$$

8. *柯西分布(Cauchy distribution)

定义（柯西分布）

若随机变量ξ具有密度函数

$$p(x)=\frac{1}{\pi}\frac{\lambda}{\lambda^2+(x-\mu)^2},\quad -\infty<x<\infty, \tag{2.2.6}$$

其中$\lambda>0$, μ为任意常数. 则称ξ服从参数为(λ,μ)的柯西分布.

$\lambda=1,\mu=0$时的柯西分布的密度函数为

$$p(x)=\frac{1}{\pi}\frac{1}{1+x^2},\quad -\infty<x<\infty.$$

容易验证(2.2.6)定义的$p(x)$满足密度函数的两个条件.

随机向量

在很多随机现象中，对同一个随机试验我们往往需要同时考察几个随机变量.

例如：发射一枚炮弹，需要同时研究炮弹弹着点的2个或3个坐标；保研排名时，需要同时参考申请保研者的数学分析、高等代数、概率论、实变函数等多门课程的成绩.

定义（n维随机向量）

若随机变量$\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega)$定义在同一概率空间$(\Omega, \mathscr{F}, P)$上，就称

$$\boldsymbol{\xi}(\omega) = (\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega))$$

为n维随机向量(n-dimensional random vector)或n维随机变量(n-dimensional random variable).

定义（n维随机向量）

若随机变量$\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega)$定义在同一概率空间$(\Omega, \mathscr{F}, P)$上，就称

$$\boldsymbol{\xi}(\omega) = (\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega))$$

为n维随机向量(n-dimensional random vector)或n维随机变量(n-dimensional random variable).

与一维情形类似，我们主要讨论离散型与连续型两种多维随机向量．同时，我们着重研究二维情形，任意n维情形可类似推广．

2.3.1 离散型随机向量

2.3.1 离散型随机向量

若随机向量只取有限组或可列无穷多组值，就称它为离散型随机向量. 对于它，我们只需列出所有各组可能值以及取这些值的概率，就可全面描述其概率分布.

2.3.1 离散型随机向量

若随机向量只取有限组或可列无穷多组值, 就称它为离散型随机向量. 对于它, 我们只需列出所有各组可能值以及取这些值的概率, 就可全面描述其概率分布.

定义 (联合分布列)

若二维随机向量(ξ,η)所有可能取值为$(x_i,y_j),i,j=1,2,\cdots$,且

$$P(\xi=x_i,\eta=y_j)=p_{ij},\ \ i,j=1,2,\cdots, \tag{2.3.1}$$

则称(2.3.1)为(ξ,η)的(联合)分布列.

(ξ, η)的联合分布列也可以用下列的表格表示:

$\xi \diagdown \eta$	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	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}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

n维离散型随机向量的联合分布列可写为

$$\mathsf{P}(\xi_1 = x_{i_1}, \cdots, \xi_n = x_{i_n}) = p_{i_1 \cdots i_n}, \quad i_1, \cdots, i_n = 1, 2, \cdots .$$

联合分布列的性质(以二维为例):

联合分布列的性质(以二维为例):

- $p_{ij} \geq 0, \ \ i, j = 1, 2, \cdots;$
- $\displaystyle\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij} = 1.$

联合分布列的性质(以二维为例):

- $p_{ij} \geq 0,\ \ i, j = 1, 2, \cdots;$
- $\displaystyle\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij} = 1.$

有了联合分布列, 任何事件的概率都可以用联合分布列算出. 例如对于任意的二维Borel集B^2,

$$\mathsf{P}((\xi, \eta) \in B^2) = \sum_{(x_i, y_j) \in B^2} p_{ij}.$$

在二维随机向量(ξ,η)中, ξ,η各作为一维随机变量有它们各自的分布列.

在二维随机向量(ξ,η)中, ξ,η各作为一维随机变量有它们各自的分布列.

对于ξ, 它只能取值$x_1,x_2,\cdots,x_i,\cdots$这些值,

在二维随机向量(ξ,η)中, ξ,η各作为一维随机变量有它们各自的分布列.

对于ξ, 它只能取值$x_1,x_2,\cdots,x_i,\cdots$这些值, 且

$$\begin{aligned}
\mathsf{P}(\xi=x_i)=&\mathsf{P}(\xi=x_i,\bigcup_{j=1}^{\infty}\{\eta=y_j\})\\
=&\mathsf{P}(\bigcup_{j=1}^{\infty}\{\xi=x_i,\eta=y_j\})\\
=&\sum_{j=1}^{\infty}\mathsf{P}(\xi=x_i,\eta=y_j)\\
=&\sum_{j=1}^{\infty}p_{ij}=:p_{i\cdot},\quad i=1,2,\cdots.
\end{aligned}\tag{2.3.2}$$

在二维随机向量(ξ,η)中, ξ,η各作为一维随机变量有它们各自的分布列.

对于ξ, 它只能取值$x_1,x_2,\cdots,x_i,\cdots$这些值, 且

$$
\begin{aligned}
\mathsf{P}(\xi=x_i)=&\mathsf{P}(\xi=x_i,\bigcup_{j=1}^{\infty}\{\eta=y_j\})\\
=&\mathsf{P}(\bigcup_{j=1}^{\infty}\{\xi=x_i,\eta=y_j\})\\
=&\sum_{j=1}^{\infty}\mathsf{P}(\xi=x_i,\eta=y_j)\\
=&\sum_{j=1}^{\infty}p_{ij}=:p_{i\cdot},\quad i=1,2,\cdots.
\end{aligned}
\tag{2.3.2}
$$

同理可得

$$\begin{aligned}\mathsf{P}(\eta = y_j) &= \sum_{i=1}^{\infty} \mathsf{P}(\xi = x_i, \eta = y_j) \\ &= \sum_{i=1}^{\infty} p_{ij} =: p_{\cdot j}, \quad j = 1, 2, \cdots . \end{aligned} \tag{2.3.3}$$

同理可得

$$\begin{aligned}\mathsf{P}(\eta = y_j) &= \sum_{i=1}^{\infty} \mathsf{P}(\xi = x_i, \eta = y_j) \\ &= \sum_{i=1}^{\infty} p_{ij} =: p_{\cdot j}, \quad j = 1, 2, \cdots . \end{aligned} \tag{2.3.3}$$

我们称(2.3.2)和(2.3.3)分别为ξ和η的边际/边缘分布列(marginal distribution).

例1 口袋中有2个白球3个黑球，连取两次，每次任取一球. 设ξ为第一次得白球数，η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形，分别求(ξ,η)的联合分布列.

例1 口袋中有2个白球3个黑球, 连取两次, 每次任取一球. 设ξ为第一次得白球数, η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形, 分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1, 两次取球结果是独立的. 因此

$$\mathsf{P}(\xi=0,\eta=0)=$$

例1 口袋中有2个白球3个黑球，连取两次，每次任取一球. 设ξ为第一次得白球数，η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形，分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1，两次取球结果是独立的. 因此

$$\mathsf{P}(\xi=0,\eta=0)=\mathsf{P}(\xi=0)\mathsf{P}(\eta=0)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25};$$

例1 口袋中有2个白球3个黑球, 连取两次, 每次任取一球. 设ξ为第一次得白球数, η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形, 分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1, 两次取球结果是独立的. 因此

$$\mathsf{P}(\xi=0,\eta=0)=\mathsf{P}(\xi=0)\mathsf{P}(\eta=0)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25};$$

$$\mathsf{P}(\xi=0,\eta=1)=$$

例1 口袋中有2个白球3个黑球, 连取两次, 每次任取一球. 设ξ为第一次得白球数, η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形, 分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1, 两次取球结果是独立的. 因此

$$\mathrm{P}(\xi=0,\eta=0)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=0)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25};$$
$$\mathrm{P}(\xi=0,\eta=1)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=1)=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25};$$

例1 口袋中有2个白球3个黑球, 连取两次, 每次任取一球. 设ξ为第一次得白球数, η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形, 分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1, 两次取球结果是独立的. 因此

$$\mathrm{P}(\xi=0,\eta=0)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=0)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25};$$
$$\mathrm{P}(\xi=0,\eta=1)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=1)=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25};$$

$$\mathrm{P}(\xi=1,\eta=0)=$$

例1 口袋中有2个白球3个黑球, 连取两次, 每次任取一球. 设ξ为第一次得白球数, η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形, 分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1, 两次取球结果是独立的. 因此

$$\mathsf{P}(\xi=0,\eta=0)=\mathsf{P}(\xi=0)\mathsf{P}(\eta=0)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25};$$
$$\mathsf{P}(\xi=0,\eta=1)=\mathsf{P}(\xi=0)\mathsf{P}(\eta=1)=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25};$$
$$\mathsf{P}(\xi=1,\eta=0)=\mathsf{P}(\xi=1)\mathsf{P}(\eta=0)=\frac{2}{5}\cdot\frac{3}{5}=\frac{6}{25};$$

例1 口袋中有2个白球3个黑球，连取两次，每次任取一球. 设ξ为第一次得白球数，η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形，分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1，两次取球结果是独立的. 因此

$$P(\xi=0,\eta=0)=P(\xi=0)P(\eta=0)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25};$$
$$P(\xi=0,\eta=1)=P(\xi=0)P(\eta=1)=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25};$$
$$P(\xi=1,\eta=0)=P(\xi=1)P(\eta=0)=\frac{2}{5}\cdot\frac{3}{5}=\frac{6}{25};$$

$$P(\xi=1,\eta=1)=$$

例1 口袋中有2个白球3个黑球, 连取两次, 每次任取一球. 设ξ为第一次得白球数, η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形, 分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1, 两次取球结果是独立的. 因此

$$
\begin{aligned}
\mathsf{P}(\xi=0,\eta=0)&=\mathsf{P}(\xi=0)\mathsf{P}(\eta=0)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25};\\
\mathsf{P}(\xi=0,\eta=1)&=\mathsf{P}(\xi=0)\mathsf{P}(\eta=1)=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25};\\
\mathsf{P}(\xi=1,\eta=0)&=\mathsf{P}(\xi=1)\mathsf{P}(\eta=0)=\frac{2}{5}\cdot\frac{3}{5}=\frac{6}{25};\\
\mathsf{P}(\xi=1,\eta=1)&=\mathsf{P}(\xi=1)\mathsf{P}(\eta=1)=\frac{2}{5}\cdot\frac{2}{5}=\frac{4}{25}.
\end{aligned}
$$

例1 口袋中有2个白球3个黑球, 连取两次, 每次任取一球. 设ξ为第一次得白球数, η为第二次得白球数. 对(1) 有放回 与 (2)不放回 两种情形, 分别求(ξ,η)的联合分布列.

解: (1) ξ和η的可能取值都是0和1, 两次取球结果是独立的. 因此

$$\mathrm{P}(\xi=0,\eta=0)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=0)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25};$$
$$\mathrm{P}(\xi=0,\eta=1)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=1)=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25};$$
$$\mathrm{P}(\xi=1,\eta=0)=\mathrm{P}(\xi=1)\mathrm{P}(\eta=0)=\frac{2}{5}\cdot\frac{3}{5}=\frac{6}{25};$$
$$\mathrm{P}(\xi=1,\eta=1)=\mathrm{P}(\xi=1)\mathrm{P}(\eta=1)=\frac{2}{5}\cdot\frac{2}{5}=\frac{4}{25}.$$

所以(ξ,η)的联合分布列可写为

$\xi\diagdown\eta$	0	1
0	9/25	6/25
1	6/25	4/25

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$\mathsf{P}(\xi = 0, \eta = 0) =$$

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$\mathsf{P}(\xi=0,\eta=0)=\mathsf{P}(\xi=0)\mathsf{P}(\eta=0|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$\mathsf{P}(\xi=0,\eta=0)=\mathsf{P}(\xi=0)\mathsf{P}(\eta=0|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$

$$\mathsf{P}(\xi=0,\eta=1)=$$

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$\mathrm{P}(\xi=0,\eta=0)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=0|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$

$$\mathrm{P}(\xi=0,\eta=1)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=1|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$P(\xi=0,\eta=0)=P(\xi=0)P(\eta=0|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$
$$P(\xi=0,\eta=1)=P(\xi=0)P(\eta=1|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$
$$P(\xi=1,\eta=0)=$$

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$P(\xi=0,\eta=0)=P(\xi=0)P(\eta=0|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$
$$P(\xi=0,\eta=1)=P(\xi=0)P(\eta=1|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$
$$P(\xi=1,\eta=0)=P(\xi=1)P(\eta=0|\xi=1)=\frac{2}{5}\cdot\frac{3}{4}=\frac{3}{10};$$

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$P(\xi=0,\eta=0)=P(\xi=0)P(\eta=0|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$

$$P(\xi=0,\eta=1)=P(\xi=0)P(\eta=1|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$

$$P(\xi=1,\eta=0)=P(\xi=1)P(\eta=0|\xi=1)=\frac{2}{5}\cdot\frac{3}{4}=\frac{3}{10};$$

$$P(\xi=1,\eta=1)=$$

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$\begin{aligned}
\mathsf{P}(\xi=0,\eta=0)=\mathsf{P}(\xi=0)\mathsf{P}(\eta=0|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};\\
\mathsf{P}(\xi=0,\eta=1)=\mathsf{P}(\xi=0)\mathsf{P}(\eta=1|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};\\
\mathsf{P}(\xi=1,\eta=0)=\mathsf{P}(\xi=1)\mathsf{P}(\eta=0|\xi=1)=\frac{2}{5}\cdot\frac{3}{4}=\frac{3}{10};\\
\mathsf{P}(\xi=1,\eta=1)=\mathsf{P}(\xi=1)\mathsf{P}(\eta=1|\xi=1)=\frac{2}{5}\cdot\frac{1}{4}=\frac{1}{10}.
\end{aligned}$$

(2) ξ和η的可能取值仍是0和1, 但两次取球结果是不独立的. 利用概率的乘法公式得:

$$\mathrm{P}(\xi=0,\eta=0)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=0|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$
$$\mathrm{P}(\xi=0,\eta=1)=\mathrm{P}(\xi=0)\mathrm{P}(\eta=1|\xi=0)=\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10};$$
$$\mathrm{P}(\xi=1,\eta=0)=\mathrm{P}(\xi=1)\mathrm{P}(\eta=0|\xi=1)=\frac{2}{5}\cdot\frac{3}{4}=\frac{3}{10};$$
$$\mathrm{P}(\xi=1,\eta=1)=\mathrm{P}(\xi=1)\mathrm{P}(\eta=1|\xi=1)=\frac{2}{5}\cdot\frac{1}{4}=\frac{1}{10}.$$

所以(ξ,η)的联合分布列可写为

$\xi\diagdown\eta$	0	1
0	$3/10$	$3/10$
1	$3/10$	$1/10$

例2 求例1的边际分布列.

例2 求例1的边际分布列.

有放回情形:

$\xi \diagdown \eta$	0	1	$p_{i\cdot}$
0	$9/25$	$6/25$	$3/5$
1	$6/25$	$4/25$	$2/5$
$p_{\cdot j}$	$3/5$	$2/5$	1

例2 求例1的边际分布列.

有放回情形:

$\xi\diagdown\eta$	0	1	$p_{i\cdot}$
0	$9/25$	$6/25$	$3/5$
1	$6/25$	$4/25$	$2/5$
$p_{\cdot j}$	$3/5$	$2/5$	1

不放回情形:

$\xi\diagdown\eta$	0	1	$p_{i\cdot}$
0	$3/10$	$3/10$	$3/5$
1	$3/10$	$1/10$	$2/5$
$p_{\cdot j}$	$3/5$	$2/5$	1

我们知道联合分布列可以确定边际分布列. 同时, 由上述两个例子可知, 给定边际分布列, 联合分布列却不能被确定, 因为还需考虑分量间的相互关系. (整体性状可以确定个体性状, 但个体性状不能确定整体性状)

2.3.2 分布函数

类似于一维随机变量, 对一般的多维随机向量, 无法用分布列来表示其概率分布, 但可以用分布函数来表示.

2.3.2 分布函数

类似于一维随机变量, 对一般的多维随机向量, 无法用分布列来表示其概率分布, 但可以用分布函数来表示. 对于任意n个实数$x_1, \cdots, x_n$, 因为$\{\omega : \xi_i(\omega) \le x_i\} \in \mathscr{F}$, 因此对于$\mathbb{R}^n$中的$n$维区间

$$c_n = \prod_{i=1}^{n} (-\infty, x_i]$$

有

$$\{\omega : \boldsymbol{\xi}(\omega) \in c_n\} = \bigcap_{i=1}^{n} \{\omega : \xi_i(\omega) \le x_i\}$$

2.3.2 分布函数

类似于一维随机变量, 对一般的多维随机向量, 无法用分布列来表示其概率分布, 但可以用分布函数来表示. 对于任意n个实数$x_1, \cdots, x_n$, 因为$\{\omega : \xi_i(\omega) \leq x_i\} \in \mathscr{F}$, 因此对于$\mathbb{R}^n$中的$n$维区间

$$c_n = \prod_{i=1}^{n}(-\infty, x_i]$$

有

$$\{\omega : \boldsymbol{\xi}(\omega) \in c_n\} = \bigcap_{i=1}^{n}\{\omega : \xi_i(\omega) \leq x_i\} \in \mathscr{F}. \tag{2.3.4}$$

2.3.2 分布函数

类似于一维随机变量, 对一般的多维随机向量, 无法用分布列来表示其概率分布, 但可以用分布函数来表示. 对于任意n个实数$x_1, \cdots, x_n$, 因为$\{\omega : \xi_i(\omega) \le x_i\} \in \mathscr{F}$, 因此对于$\mathbb{R}^n$中的$n$维区间

$$c_n = \prod_{i=1}^{n} (-\infty, x_i]$$

有

$$\{\omega : \boldsymbol{\xi}(\omega) \in c_n\} = \bigcap_{i=1}^{n} \{\omega : \xi_i(\omega) \le x_i\} \in \mathscr{F}. \tag{2.3.4}$$

进一步还可以证明: 对$\mathbb{R}^n$上任一Borel集B^n, $\{\boldsymbol{\xi}(\omega) \in B^n\}$ 的概率都可以通过(2.3.4)的概率表示. 因此可以用(2.3.4)的概率代表$\boldsymbol{\xi}(\omega)$的概率分布.

定义（联合分布函数）

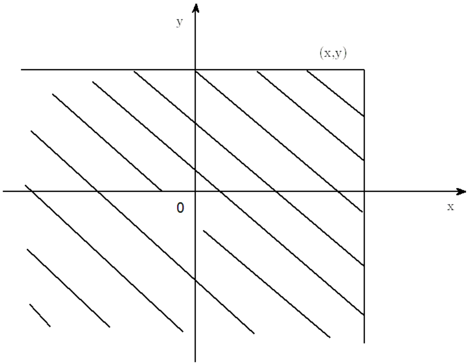
对任意的$(x_1,\cdots,x_n)\in\mathbb{R}^n$, 称$n$元函数

$$F(x_1,\cdots,x_n)=P(\xi_1(\omega)\le x_1,\cdots,\xi_n(\omega)\le x_n)$$

为随机向量$\boldsymbol{\xi}(\omega)=(\xi_1(\omega),\cdots,\xi_n(\omega))$的(联合)分布函数((joint) distribution function).

二维情形的分布函数:

$$F(x,y) = \mathsf{P}(\xi \le x, \eta \le y).$$



对于矩形区域

$$I: \; a_1 < x \le b_1, \; a_2 < y \le b_2,$$

易知

$\mathsf{P}((\xi, \eta) \in I) =$

对于矩形区域

$$I:\ a_1 < x \leq b_1,\ a_2 < y \leq b_2,$$

易知

$$\mathsf{P}((\xi,\eta)\in I) = F(b_1,b_2) - F(a_1,b_2) - F(b_1,a_2) + F(a_1,a_2).$$

二维联合分布函数$F(x,y)$的性质:

二维联合分布函数$F(x,y)$的性质:

- $F(x,y)$对每个变量单调不减;

二维联合分布函数$F(x,y)$的性质:

- $F(x,y)$对每个变量单调不减;
- 对任意的(x,y),

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

二维联合分布函数$F(x,y)$的性质:

- $F(x,y)$对每个变量单调不减;
- 对任意的(x,y),

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

- $F(x,y)$对每个变量右连续;

二维联合分布函数$F(x,y)$的性质:

- $F(x,y)$对每个变量单调不减;
- 对任意的(x,y),

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

- $F(x,y)$对每个变量右连续;

二维联合分布函数$F(x,y)$的性质:

- $F(x,y)$对每个变量单调不减;
- 对任意的(x,y),

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

- $F(x,y)$对每个变量右连续;
- 对任意的$a_1<b_1,\ a_2<b_2$,

$$F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)+F(a_1,a_2)\geq 0.$$

ξ, η作为随机变量, 有自己的分布函数. 他们的分布函数被称为边际/边缘分布函数.

联合分布函数与边际分布函数的关系:

ξ, η作为随机变量, 有自己的分布函数. 他们的分布函数被称为边际/边缘分布函数.

联合分布函数与边际分布函数的关系:

$$F_\xi(x) = \mathsf{P}(\xi \leq x) =$$

ξ, η作为随机变量, 有自己的分布函数. 他们的分布函数被称为边际/边缘分布函数.

联合分布函数与边际分布函数的关系:

$$F_\xi(x) = \mathsf{P}(\xi \leq x) = \mathsf{P}(\xi \leq x, \eta < \infty) = F(x, \infty),$$
$$F_\eta(y) = \mathsf{P}(\eta \leq y) =$$

ξ,η作为随机变量, 有自己的分布函数. 他们的分布函数被称为边际/边缘分布函数.

联合分布函数与边际分布函数的关系:

$$F_\xi(x) = \mathsf{P}(\xi \le x) = \mathsf{P}(\xi \le x, \eta < \infty) = F(x, \infty),$$
$$F_\eta(y) = \mathsf{P}(\eta \le y) = \mathsf{P}(\xi < \infty, \eta \le y) = F(\infty, y).$$

2.3.3 连续型随机向量(continuous random vector)

2.3.3 连续型随机向量(continuous random vector)

我们来给出连续型随机向量及其联合概率密度函数的定义.

2.3.3 连续型随机向量(continuous random vector)

我们来给出连续型随机向量及其联合概率密度函数的定义.

定义 (连续型随机向量)

若存在n元可积的非负函数$p(x_1,\cdots,x_n)$, 使得n元分布函数可表示为

$$F(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}p(y_1,y_2,\cdots,y_n)\mathrm{d}y_1\cdots\mathrm{d}y_n,$$

则称所对应的随机向量为n维连续型随机向量, 称$p(y_1,\cdots,y_n)$为相应的(联合)概率密度函数, 简称(联合)密度函数.

联合密度函数$p(y_1, y_2, \cdots, y_n)$的性质:

联合密度函数$p(y_1, y_2, \cdots, y_n)$的性质:

- $p(y_1, \cdots, y_n) \geq 0$;

联合密度函数$p(y_1, y_2, \cdots, y_n)$的性质:

- $p(y_1, \cdots, y_n) \geq 0$;
- $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(y_1, \cdots, y_n) \mathrm{d}y_1 \cdots \mathrm{d}y_n = 1$.

联合密度函数$p(y_1, y_2, \cdots, y_n)$的性质:

- $p(y_1, \cdots, y_n) \geq 0$;
- $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(y_1, \cdots, y_n) \mathrm{d}y_1 \cdots \mathrm{d}y_n = 1.$

注: 与一维情形类似, 除了一个零测度集外, 可以对(联合)分布函数求导得到(联合)密度函数:

$$\frac{\partial^n F(x_1, \cdots, x_n)}{\partial x_1 \cdots \partial x_n} = p(x_1, \cdots, x_n).$$

同时, 对任一n维Borel集$B^n \in \mathscr{B}^n$ (n维Borel域), 有

$$\begin{aligned} & \mathsf{P}(\boldsymbol{\xi}(\omega) \in B^n) \\ = & \int \cdots \int_{(x_1, \cdots, x_n) \in B^n} p(x_1, \cdots, x_n) \mathrm{d}x_1 \cdots \mathrm{d}x_n. \end{aligned} \tag{2.3.5}$$

ξ_i自己也有密度函数, 它们的密度函数被称为边际/边缘(概率)密度函数.

ξ_i自己也有密度函数, 它们的密度函数被称为边际/边缘(概率)密度函数. 联合密度函数与边际密度函数的关系, 以二维为例:

ξ_i自己也有密度函数，它们的密度函数被称为边际/边缘(概率)密度函数. 联合密度函数与边际密度函数的关系，以二维为例:

设(ξ, η)的密度函数为$p(x, y)$，分布函数为$F(x, y)$. 则ξ的边际分布函数为

$$
\begin{aligned}
F_\xi(x) =& F(x, \infty) \\
=& \int_{-\infty}^{x} \int_{-\infty}^{\infty} p(u, v) \mathrm{d}u \mathrm{d}v \\
=& \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} p(u, v) \mathrm{d}v \right) \mathrm{d}u.
\end{aligned}
$$

所以由一维连续型随机变量及其密度函数的定义可知ξ的边际密度函数为

$$
p_\xi(u) = \int_{-\infty}^{\infty} p(u, v) \mathrm{d}v.
$$

ξ_i自己也有密度函数, 它们的密度函数被称为边际/边缘(概率)密度函数. 联合密度函数与边际密度函数的关系, 以二维为例:

设(ξ,η)的密度函数为$p(x,y)$, 分布函数为$F(x,y)$. 则ξ的边际分布函数为

$$\begin{aligned} F_\xi(x) =& F(x,\infty) \\ =& \int_{-\infty}^{x}\int_{-\infty}^{\infty} p(u,v)\mathrm{d}u\mathrm{d}v \\ =& \int_{-\infty}^{x}\Big(\int_{-\infty}^{\infty} p(u,v)\mathrm{d}v\Big)\mathrm{d}u. \end{aligned}$$

ξ_i自己也有密度函数, 它们的密度函数被称为边际/边缘(概率)密度函数. 联合密度函数与边际密度函数的关系, 以二维为例:

设(ξ, η)的密度函数为$p(x, y)$, 分布函数为$F(x, y)$. 则ξ的边际分布函数为

$$\begin{aligned} F_\xi(x) =& F(x, \infty) \\ =& \int_{-\infty}^{x} \int_{-\infty}^{\infty} p(u, v) \mathrm{d}u \mathrm{d}v \\ =& \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} p(u, v) \mathrm{d}v \right) \mathrm{d}u. \end{aligned}$$

所以由一维连续型随机变量及其密度函数的定义可知ξ的边际密度函数为

$$p_\xi(u) = \int_{-\infty}^{\infty} p(u, v) \mathrm{d}v.$$

同理, 可得

$$p_\eta(v) = \int_{-\infty}^{\infty} p(u, v)\mathrm{d}u.$$

例3 设二维随机向量(ξ,η)的密度函数为

$$p(x,y)=\begin{cases} Ae^{-2(x+y)}, & x>0, y>0, \\ 0, & \text{其他}. \end{cases}$$

(1) 确定常数A; (2) 求分布函数; (3) 求边际密度函数; (4) 计算概率P$(\xi<1,\eta<2)$; (5) 计算概率P$(\xi+\eta<1)$.

例3 设二维随机向量(ξ,η)的密度函数为

$$p(x,y)=\begin{cases} Ae^{-2(x+y)}, & x>0, y>0, \\ 0, & \text{其他}. \end{cases}$$

(1) 确定常数A; (2) 求分布函数; (3) 求边际密度函数; (4) 计算概率P$(\xi<1,\eta<2)$; (5) 计算概率P$(\xi+\eta<1)$.

解: (1) 利用联合密度函数的性质, 有

$$1=\int_0^{\infty}\int_0^{\infty} Ae^{-2(x+y)}\mathrm{d}x\mathrm{d}y=$$

例3 设二维随机向量(ξ,η)的密度函数为

$$p(x,y)=\begin{cases} Ae^{-2(x+y)}, & x>0, y>0, \\ 0, & \text{其他}. \end{cases}$$

(1) 确定常数A; (2) 求分布函数; (3) 求边际密度函数; (4) 计算概率P$(\xi<1,\eta<2)$; (5) 计算概率P$(\xi+\eta<1)$.

解: (1) 利用联合密度函数的性质, 有

$$1=\int_0^{\infty}\int_0^{\infty} Ae^{-2(x+y)}\mathrm{d}x\mathrm{d}y=\frac{A}{4},$$

所以$A=4$.

(2) 我们根据公式

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

来计算分布函数.

(2) 我们根据公式

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

来计算分布函数. 当$x\le 0$或者$y\le 0$时, $p(x,y)=0$. 因此$F(x,y)=0$;

(2) 我们根据公式

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

来计算分布函数. 当$x\le 0$或者$y\le 0$时, $p(x,y)=0$. 因此$F(x,y)=0$; 当$x>0$且$y>0$时,

$$\begin{aligned}F(x,y)=&\int_{0}^{x}\int_{0}^{y}4e^{-2(u+v)}\mathrm{d}u\mathrm{d}v\\=&(1-e^{-2x})(1-e^{-2y}).\end{aligned}$$

(2) 我们根据公式

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

来计算分布函数. 当$x\leq 0$或者$y\leq 0$时, $p(x,y)=0$. 因此$F(x,y)=0$; 当$x>0$且$y>0$时,

$$\begin{aligned}F(x,y)=&\int_{0}^{x}\int_{0}^{y}4e^{-2(u+v)}\mathrm{d}u\mathrm{d}v\\=&(1-e^{-2x})(1-e^{-2y}).\end{aligned}$$

所以(ξ,η)的分布函数为

$$F(x,y)=\begin{cases}(1-e^{-2x})(1-e^{-2y}), & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

(3) ξ的边际密度函数为

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y.$$

(3) ξ的边际密度函数为

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y.$$

当$x \leq 0$时, $p(x,y) = 0$. 因此$p_\xi(x) = 0$;

(3) ξ的边际密度函数为

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y.$$

当$x \le 0$时, $p(x,y) = 0$. 因此$p_\xi(x) = 0$; 当$x > 0$时,

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y = \int_{0}^{\infty} 4e^{-2(x+y)}\mathrm{d}y = 2e^{-2x}.$$

(3) ξ的边际密度函数为

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y.$$

当$x \le 0$时, $p(x,y) = 0$. 因此$p_\xi(x) = 0$; 当$x > 0$时,

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y = \int_0^{\infty} 4e^{-2(x+y)}\mathrm{d}y = 2e^{-2x}.$$

因此, ξ的边际密度函数为

$$p_\xi(x) = \begin{cases} 2e^{-2x}, & x > 0, \\ 0, & \text{其他}. \end{cases}$$

(3) ξ的边际密度函数为

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y.$$

当$x \le 0$时, $p(x,y) = 0$. 因此$p_\xi(x) = 0$; 当$x > 0$时,

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y = \int_0^{\infty} 4e^{-2(x+y)}\mathrm{d}y = 2e^{-2x}.$$

因此, ξ的边际密度函数为

$$p_\xi(x) = \begin{cases} 2e^{-2x}, & x > 0, \\ 0, & \text{其他}. \end{cases}$$

同理可得η的边际密度函数为

$$p_\eta(x) = \begin{cases} 2e^{-2y}, & y > 0, \\ 0, & \text{其他}. \end{cases}$$

(4) $\mathrm{P}(\xi < 1, \eta < 2) =$

(4) $\mathrm{P}(\xi < 1, \eta < 2) = F(1,2) = (1-e^{-2})(1-e^{-4})$.

(4) $\mathrm{P}(\xi < 1, \eta < 2) = F(1,2) = (1-e^{-2})(1-e^{-4})$.

(5) 根据公式(2.3.5)可得:

$$\begin{aligned}\mathrm{P}(\xi+\eta<1) =& \iint_{\substack{x+y<1 \\ x>0,y>0}} 4e^{-2(x+y)}\mathrm{d}x\mathrm{d}y \\ =& \int_0^1 \Big(\int_0^{1-x} 4e^{-2(x+y)}\mathrm{d}y \Big)\mathrm{d}x \\ =& 1-3e^{-2}.\end{aligned}$$

(4) $\mathrm{P}(\xi < 1, \eta < 2) = F(1,2) = (1-e^{-2})(1-e^{-4})$.

(5) 根据公式(2.3.5)可得:

$$
\begin{aligned}
\mathrm{P}(\xi+\eta<1) =& \int\int_{\substack{x+y<1 \\ x>0,y>0}} 4e^{-2(x+y)}\mathrm{d}x\mathrm{d}y \\
=& \int_0^1 \Big(\int_0^{1-x} 4e^{-2(x+y)}\mathrm{d}y\Big)\mathrm{d}x \\
=& 1-3e^{-2}.
\end{aligned}
$$

两个重要的连续型随机向量:

两个重要的连续型随机向量:

1. n维均匀分布

两个重要的连续型随机向量:

1. n维均匀分布

定义 (n维均匀分布)

若n维向量$\boldsymbol{\xi}$具有密度函数

$$p(x_1, \cdots, x_n) = \begin{cases} 1/S_G, & (x_1, \cdots, x_n) \in G, \\ 0, & \text{其他}. \end{cases}$$

其中, G是R^n中的一个*Borel*集, S_G为G的测度(二维时, S_G表示G的面积; 三维时, S_G表示G的体积), 则称$\boldsymbol{\xi}$服从G上的均匀分布.

例4 (ξ, η)在圆形区域$x^2 + y^2 \le 1$上服从均匀分布，求ξ, η的边际密度函数.

例4 (ξ,η)在圆形区域$x^2+y^2\leq 1$上服从均匀分布，求ξ,η的边际密度函数.

解: 联合密度函数为

$$p(x,y)=\begin{cases} 1/\pi, & x^2+y^2\leq 1, \\ 0, & \text{其他}. \end{cases}$$

例4 (ξ, η)在圆形区域$x^2+y^2\leq 1$上服从均匀分布，求ξ, η的边际密度函数.

解：联合密度函数为

$$p(x,y)=\begin{cases} 1/\pi, & x^2+y^2\leq 1, \\ 0, & \text{其他}. \end{cases}$$

当$|x|>1$时，$p(x,y)=0$. 所以此时$p_\xi(x)=0$.

例4 (ξ,η)在圆形区域$x^2+y^2\le 1$上服从均匀分布, 求ξ,η的边际密度函数.

解: 联合密度函数为

$$p(x,y)=\begin{cases}1/\pi, & x^2+y^2\le 1,\\ 0, & \text{其他}.\end{cases}$$

当$|x|>1$时, $p(x,y)=0$. 所以此时$p_\xi(x)=0$. 当$|x|\le 1$时

$$p_\xi(x)=\int_{-\infty}^{\infty}p(x,y)\mathrm{d}y=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}\mathrm{d}y=\frac{2}{\pi}\sqrt{1-x^2}.$$

例4 (ξ,η)在圆形区域$x^2+y^2\le 1$上服从均匀分布，求ξ,η的边际密度函数.

解：联合密度函数为

$$p(x,y)=\begin{cases} 1/\pi, & x^2+y^2\le 1, \\ 0, & \text{其他}. \end{cases}$$

当$|x|>1$时，$p(x,y)=0$. 所以此时$p_\xi(x)=0$. 当$|x|\le 1$时

$$p_\xi(x)=\int_{-\infty}^{\infty}p(x,y)\mathrm{d}y=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}\mathrm{d}y=\frac{2}{\pi}\sqrt{1-x^2}.$$

综上所述，ξ的边际密度函数

$$p_\xi(x)=\begin{cases} \frac{2}{\pi}\sqrt{1-x^2}, & |x|\le 1, \\ 0, & |x|>1. \end{cases}$$

例4 (ξ,η)在圆形区域$x^2+y^2\le 1$上服从均匀分布，求ξ,η的边际密度函数.

解：联合密度函数为

$$p(x,y)=\begin{cases}1/\pi, & x^2+y^2\le 1,\\ 0, & \text{其他}.\end{cases}$$

当$|x|>1$时，$p(x,y)=0$. 所以此时$p_\xi(x)=0$. 当$|x|\le 1$时

$$p_\xi(x)=\int_{-\infty}^{\infty}p(x,y)\mathrm{d}y=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}\mathrm{d}y=\frac{2}{\pi}\sqrt{1-x^2}.$$

综上所述，ξ的边际密度函数

$$p_\xi(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2}, & |x|\le 1,\\ 0, & |x|>1.\end{cases}$$

同理可得η的边际密度函数

$$p_\eta(y)=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2}, & |y|\le 1,\\ 0, & |y|>1.\end{cases}$$

2. n维正态分布

2. n维正态分布

一些记号: 设$\boldsymbol{\Sigma} = (\sigma_{ij})$为$n$阶正定对称矩阵, $|\boldsymbol{\Sigma}|$为其行列式, $\boldsymbol{\Sigma}^{-1}$为其逆. 又记

$$\boldsymbol{x} = (x_1, \cdots, x_n)', \quad \boldsymbol{a} = (a_1, \cdots, a_n)'.$$

2. n维正态分布

一些记号: 设$\boldsymbol{\Sigma} = (\sigma_{ij})$为$n$阶正定对称矩阵, $|\boldsymbol{\Sigma}|$为其行列式, $\boldsymbol{\Sigma}^{-1}$为其逆. 又记

$$\boldsymbol{x} = (x_1, \cdots, x_n)', \quad \boldsymbol{a} = (a_1, \cdots, a_n)'.$$

定义 (n维正态分布)

若n维随机向量$\boldsymbol{\xi}$具有密度函数

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{a})\right\}, \quad \boldsymbol{x} \in \mathbb{R}^n. \tag{2.3.6}$$

则称$\boldsymbol{\xi}$服从n维正态分布, 记作$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$.

我们先来证明(2.3.6)是一个密度函数. 首先, $p(\boldsymbol{x}) \geq 0$是显然的. 然后我们来证明

$$\int_{\mathbb{R}^n} p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = 1.$$

我们先来证明(2.3.6)是一个密度函数. 首先, $p(\boldsymbol{x}) \geq 0$是显然的. 然后我们来证明

$$\int_{\mathbb{R}^n} p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1.$$

先考虑特殊情形: $\boldsymbol{a} = \mathbf{0}, \boldsymbol{\Sigma} = \boldsymbol{I}$, 其中$\boldsymbol{I}$是$n$阶单位阵. 此时,

$$p(\boldsymbol{x}) =$$

我们先来证明(2.3.6)是一个密度函数. 首先, $p(\boldsymbol{x}) \geq 0$是显然的. 然后我们来证明

$$\int_{\mathbb{R}^n} p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1.$$

先考虑特殊情形: $\boldsymbol{a} = \mathbf{0}, \boldsymbol{\Sigma} = \boldsymbol{I}$, 其中$\boldsymbol{I}$是$n$阶单位阵. 此时,

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{\boldsymbol{x}'\boldsymbol{x}}{2} \right\} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}, \quad \boldsymbol{x} \in \mathbb{R}^n. \tag{2.3.7}$$

所以

$$\int_{\mathbb{R}^n} p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} =$$

我们先来证明(2.3.6)是一个密度函数. 首先, $p(\boldsymbol{x}) \geq 0$是显然的. 然后我们来证明

$$\int_{\mathbb{R}^n} p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = 1.$$

先考虑特殊情形: $\boldsymbol{a} = \boldsymbol{0}, \boldsymbol{\Sigma} = \boldsymbol{I}$, 其中$\boldsymbol{I}$是$n$阶单位阵. 此时,

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{\boldsymbol{x}'\boldsymbol{x}}{2} \right\} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}, \quad \boldsymbol{x} \in \mathbb{R}^n. \text{(2.3.7)}$$

所以

$$\int_{\mathbb{R}^n} p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} \mathrm{d}x_i = 1.$$

我们先来证明(2.3.6)是一个密度函数. 首先, $p(\boldsymbol{x}) \geq 0$是显然的. 然后我们来证明

$$\int_{\mathbb{R}^n} p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = 1.$$

先考虑特殊情形: $\boldsymbol{a} = \mathbf{0}, \boldsymbol{\Sigma} = \boldsymbol{I}$, 其中$\boldsymbol{I}$是$n$阶单位阵. 此时,

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{\boldsymbol{x}'\boldsymbol{x}}{2} \right\} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}, \quad \boldsymbol{x} \in \mathbb{R}^n. \tag{2.3.7}$$

所以

$$\int_{\mathbb{R}^n} p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} \mathrm{d}x_i = 1.$$

(注: 在上述特殊情形下, 我们称(2.3.7)为n维标准正态分布的(联合)密度函数)

对于一般情形, 因为$\boldsymbol{\Sigma}$是对称正定矩阵, 所以存在$n\times n$对称正定矩阵$\boldsymbol{L}$使得

$$\boldsymbol{\Sigma}=\boldsymbol{L}\boldsymbol{L}',\quad |\boldsymbol{L}|=|\boldsymbol{\Sigma}|^{1/2}.$$

然后令$\boldsymbol{y}=\boldsymbol{L}^{-1}(\boldsymbol{x}-\boldsymbol{a})$, 可知

$$\begin{aligned}
&\int_{\mathbb{R}^n}p(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\
=&\int_{\mathbb{R}^n}\frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\right\}|\boldsymbol{L}|\mathrm{d}\boldsymbol{y}\\
=&\int_{\mathbb{R}^n}\frac{1}{(2\pi)^{n/2}}\exp\left\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\right\}\mathrm{d}\boldsymbol{y}\\
=&\prod_{i=1}^{n}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y_i^2}{2}}\mathrm{d}y_i\\
=&1.
\end{aligned}$$

对于一般情形，因为$\boldsymbol{\Sigma}$是对称正定矩阵，所以存在$n \times n$对称正定矩阵$\boldsymbol{L}$使得

$$\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}', \quad |\boldsymbol{L}| = |\boldsymbol{\Sigma}|^{1/2}.$$

然后令$\boldsymbol{y} = \boldsymbol{L}^{-1}(\boldsymbol{x} - \boldsymbol{a})$，可知

$$\begin{aligned}
&\int_{\mathbb{R}^n} p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
=&\int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\right\}|\boldsymbol{L}|\mathrm{d}\boldsymbol{y} \\
=&\int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\right\}\mathrm{d}\boldsymbol{y} \\
=&\prod_{i=1}^{n}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y_i^2}{2}}\mathrm{d}y_i \\
=&1.
\end{aligned}$$

$n=1$时, $\boldsymbol{\Sigma}=\sigma^2, \boldsymbol{x}=x, \boldsymbol{a}=a$. (2.3.6)变成

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}, \quad -\infty<x<\infty.$$

$n=2$时，记

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 r \\ \sigma_1\sigma_2 r & \sigma_2^2 \end{pmatrix},$$

其中$\sigma_1 > 0, \sigma_2 > 0, |r| < 1$. 再记

$$\boldsymbol{x} = (x, y)', \quad \boldsymbol{a} = (a, b)'.$$

容易求得

$$\boldsymbol{\Sigma}^{-1} =$$

$n=2$时，记

$$\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2 r\\ \sigma_1\sigma_2 r & \sigma_2^2\end{pmatrix},$$

其中$\sigma_1>0,\sigma_2>0,|r|<1$. 再记

$$\boldsymbol{x}=(x,y)',\ \ \boldsymbol{a}=(a,b)'.$$

容易求得

$$\boldsymbol{\Sigma}^{-1}=\frac{1}{|\boldsymbol{\Sigma}|}\begin{pmatrix}\sigma_2^2 & -\sigma_1\sigma_2 r\\ -\sigma_1\sigma_2 r & \sigma_1^2\end{pmatrix}.$$

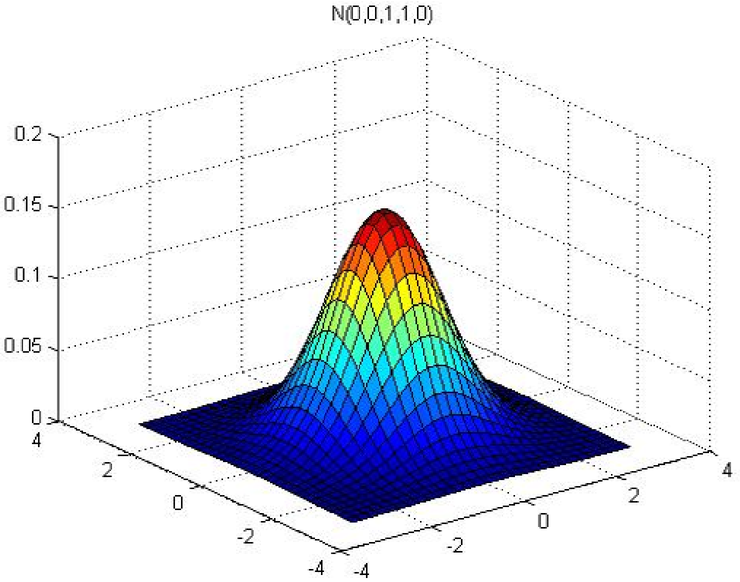
此时, (2.3.6)变成

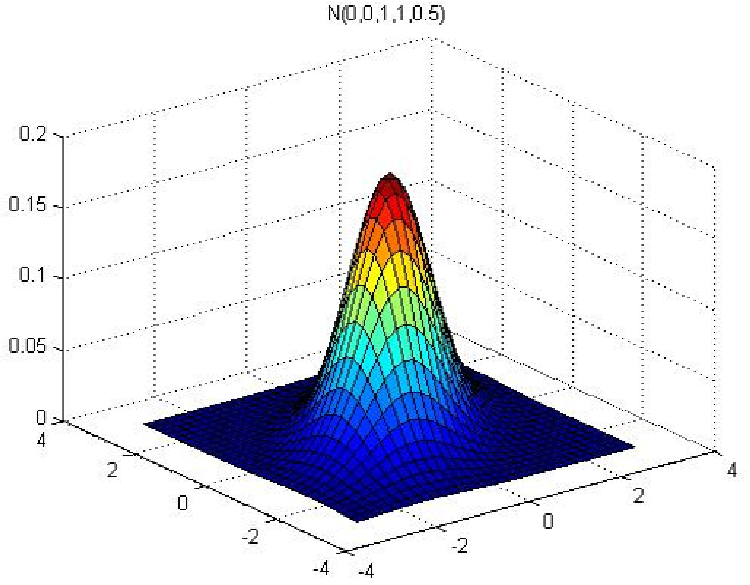
$$p(x,y) =$$

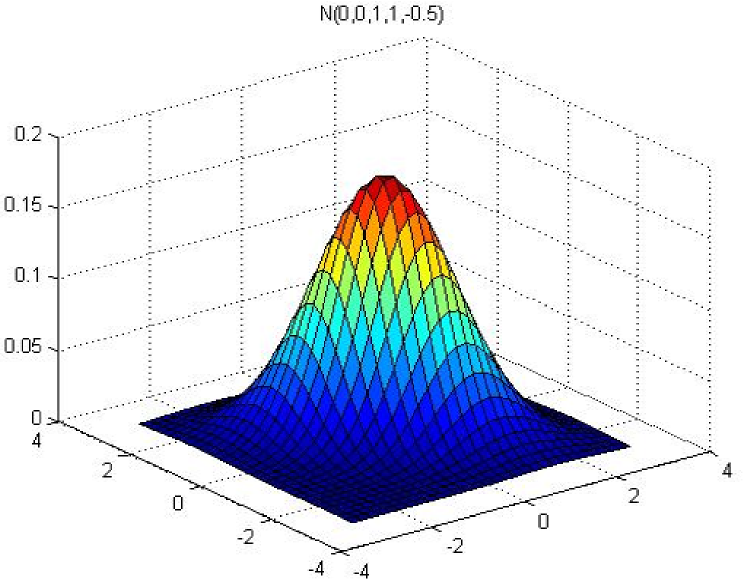
此时, (2.3.6)变成

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\cdot \exp\Big\{-\frac{1}{2(1-r^2)}\Big[\frac{(x-a)^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\Big]\Big\},\quad -\infty<x,y<\infty, \tag{2.3.8}$$

记为$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$.


N(0,0,1,1,0)
0.2
0.15
0.1
0.05
0
4
2
0
-2
-4
-4
-2
0
2
4


N(0,0,1,1,0.5)
0.2
0.15
0.1
0.05
0
4
2
0
-2
-4
-4
-2
0
2
4


N(0,0,1,1,-0.5)
0.2
0.15
0.1
0.05
0
4
2
0
-2
-4
-4
-2
0
2
4

(2.3.8)又可写为

$$
\begin{aligned}
p(x,y) =& \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{ -\frac{(x-a)^2}{2\sigma_1^2} \right\}. \\
& \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} \exp\left\{ -\frac{[y-b-\frac{r\sigma_2}{\sigma_1}(x-a)]^2}{2\sigma_2^2(1-r^2)} \right\} \\
:=& \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{ -\frac{(x-a)^2}{2\sigma_1^2} \right\} \cdot p_{\eta|\xi}(y|x).
\end{aligned}
$$

(2.3.8)又可写为

$$
\begin{aligned}
p(x,y) =& \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-a)^2}{2\sigma_1^2}\right\}.\\
&\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{\left[y-b-\frac{r\sigma_2}{\sigma_1}(x-a)\right]^2}{2\sigma_2^2(1-r^2)}\right\}\\
:=&\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-a)^2}{2\sigma_1^2}\right\}\cdot p_{\eta|\xi}(y|x).
\end{aligned}
$$

所以, ξ的边际密度函数为

$$
\begin{aligned}
p_\xi(x) =& \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y\\
=&\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-a)^2}{2\sigma_1^2}\right\}\cdot\int_{-\infty}^{\infty} p_{\eta|\xi}(y|x)\mathrm{d}y\\
=&\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-a)^2}{2\sigma_1^2}\right\},\quad x\in\mathbb{R}.
\end{aligned}
$$

所以,

$$\xi \sim N(a, \sigma_1^2).$$

所以,

$$\xi \sim N(a, \sigma_1^2).$$

同理,

$$\eta \sim N(b, \sigma_2^2).$$

所以,

$$\xi \sim N(a, \sigma_1^2).$$

同理,

$$\eta \sim N(b, \sigma_2^2).$$

上述结论说明: 二维正态分布的边际分布仍然是正态分布, 且与r无关. 但反过来不一定正确, 即(ξ, η)的边际分布都是正态分布, 其联合分布未必是二维正态分布. 举例如下:

例5 (ξ, η)的联合密度函数为

$$p(x, y) = \frac{1}{2\pi} \exp\left\{ -\frac{x^2 + y^2}{2} \right\} \cdot (1 + \sin x \cdot \sin y), \quad (x, y) \in \mathbb{R}^2.$$

求边际分布.

例5 (ξ,η)的联合密度函数为

$$p(x,y)=\frac{1}{2\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\}\cdot(1+\sin x\cdot\sin y),\quad (x,y)\in\mathbb{R}^2.$$

求边际分布.

解: ξ的密度函数为

$$\begin{aligned}
p_\xi(x) &= \int_{-\infty}^{\infty} p(x,y)\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\mathrm{d}y+\frac{1}{2\pi}e^{-\frac{x^2}{2}}\sin x\cdot\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}\sin y\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\quad x\in\mathbb{R}.
\end{aligned}$$

例5 (ξ,η)的联合密度函数为

$$p(x,y)=\frac{1}{2\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\}\cdot(1+\sin x\cdot\sin y),\quad (x,y)\in\mathbb{R}^2.$$

求边际分布.

解: ξ的密度函数为

$$\begin{aligned}p_\xi(x)&=\int_{-\infty}^{\infty}p(x,y)\mathrm{d}y\\&=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\mathrm{d}y+\frac{1}{2\pi}e^{-\frac{x^2}{2}}\sin x\cdot\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}\sin y\mathrm{d}y\\&=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\quad x\in\mathbb{R}.\end{aligned}$$

同理

$$p_\eta(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}},\quad y\in\mathbb{R}.$$

随机变量的独立性

在第一章中, 我们曾经定义过事件的独立性. 在这里, 我们将把事件的独立性的概念移植到随机变量.

随机变量的独立性

在第一章中，我们曾经定义过事件的独立性. 在这里，我们将把事件的独立性的概念移植到随机变量.

自然地，我们把ξ与η的相互独立定义为对一切$x, y \in \mathbb{R}$，事件$\{\xi \leq x\}$与$\{\eta \leq y\}$都相互独立.

定义 (随机变量的独立性)

设ξ与η为定义在同一概率空间上的随机变量. 若对一切$x, y \in \mathbb{R}$, 都有

$$P(\xi \leq x, \eta \leq y) = P(\xi \leq x)P(\eta \leq y),$$

即

$$F(x, y) = F_\xi(x)F_\eta(y), \quad \forall x, y \in \mathbb{R}. \tag{2.4.1}$$

则称随机变量ξ与η相互独立(independent). 否则, 称随机变量ξ与η相依(dependent).

若(ξ,η)为二维离散型随机向量，则(2.4.1)等价于

$$\mathsf{P}(\xi=x_i,\eta=y_j)=\mathsf{P}(\xi=x_i)\mathsf{P}(\eta=y_j),\quad \forall i,j=1,2,\cdots.$$

即等价于

$$p_{ij}=p_{i\cdot}p_{\cdot j},\quad \forall i,j=1,2,\cdots. \tag{2.4.2}$$

若(ξ,η)为二维离散型随机向量, 则(2.4.1)等价于

$$\mathrm{P}(\xi=x_i,\eta=y_j)=\mathrm{P}(\xi=x_i)\mathrm{P}(\eta=y_j),\ \ \forall i,j=1,2,\cdots.$$

即等价于

$$p_{ij}=p_{i\cdot}p_{\cdot j},\ \ \forall i,j=1,2,\cdots. \tag{2.4.2}$$

证明: (2.4.2)⇒(2.4.1):

若(ξ,η)为二维离散型随机向量, 则(2.4.1)等价于

$$\mathsf{P}(\xi=x_i,\eta=y_j)=\mathsf{P}(\xi=x_i)\mathsf{P}(\eta=y_j),\ \ \forall i,j=1,2,\cdots.$$

即等价于

$$p_{ij}=p_{i\cdot}p_{\cdot j},\ \ \forall i,j=1,2,\cdots. \tag{2.4.2}$$

证明: (2.4.2)⇒(2.4.1): 对任意的$x,y\in\mathbb{R}$,

$$\begin{aligned}\mathsf{P}(\xi\le x,\eta\le y)=&\sum_{x_i\le x}\sum_{y_j\le y}\mathsf{P}(\xi=x_i,\eta=y_j)\\ =&\sum_{x_i\le x}\mathsf{P}(\xi=x_i)\sum_{y_j\le y}\mathsf{P}(\eta=y_j)\\ =&\mathsf{P}(\xi\le x)\mathsf{P}(\eta\le y),\end{aligned}$$

即$F(x,y)=F_\xi(x)F_\eta(y)$成立.

(2.4.1)$\Rightarrow$(2.4.2):

(2.4.1)⇒(2.4.2): 不妨假设$x_i, i \geq 1$以及$y_j, j \geq 1$是按从小到大排序的, 且记$x_0 = y_0 = -\infty$. 则

$$
\begin{aligned}
&\mathsf{P}(\xi = x_i, \eta = y_j) \\
=&\mathsf{P}(x_{i-1} < \xi \leq x_i, y_{j-1} < \eta \leq y_j) \\
=&F(x_i, y_j) - F(x_i, y_{j-1}) - F(x_{i-1}, y_j) + F(x_{i-1}, y_{j-1}) \\
=&F_\xi(x_i)F_\eta(y_j) - F_\xi(x_i)F_\eta(y_{j-1}) - F_\xi(x_{i-1})F_\eta(y_j) \\
&+ F_\xi(x_{i-1})F_\eta(y_{j-1}) \\
=&F_\xi(x_i)\big[F_\eta(y_j) - F_\eta(y_{j-1})\big] - F_\xi(x_{i-1})\big[F_\eta(y_j) - F_\eta(y_{j-1})\big] \\
=&\big[F_\xi(x_i) - F_\xi(x_{i-1})\big] \cdot \big[F_\eta(y_j) - F_\eta(y_{j-1})\big] \\
=&\mathsf{P}(\xi = x_i)\mathsf{P}(\eta = y_j), \quad \forall i, j = 1, 2, \cdots.
\end{aligned}
$$

得证.

(2.4.1)⇒(2.4.2): 不妨假设$x_i, i \geq 1$以及$y_j, j \geq 1$是按从小到大排序的, 且记$x_0 = y_0 = -\infty$. 则

$$
\begin{aligned}
&\mathsf{P}(\xi = x_i, \eta = y_j) \\
=&\mathsf{P}(x_{i-1} < \xi \leq x_i, y_{j-1} < \eta \leq y_j) \\
=&F(x_i, y_j) - F(x_i, y_{j-1}) - F(x_{i-1}, y_j) + F(x_{i-1}, y_{j-1}) \\
=&F_\xi(x_i)F_\eta(y_j) - F_\xi(x_i)F_\eta(y_{j-1}) - F_\xi(x_{i-1})F_\eta(y_j) \\
&+ F_\xi(x_{i-1})F_\eta(y_{j-1}) \\
=&F_\xi(x_i)\left[F_\eta(y_j) - F_\eta(y_{j-1})\right] - F_\xi(x_{i-1})\left[F_\eta(y_j) - F_\eta(y_{j-1})\right] \\
=&\left[F_\xi(x_i) - F_\xi(x_{i-1})\right] \cdot \left[F_\eta(y_j) - F_\eta(y_{j-1})\right] \\
=&\mathsf{P}(\xi = x_i)\mathsf{P}(\eta = y_j), \quad \forall i, j = 1, 2, \cdots .
\end{aligned}
$$

得证.

若(ξ,η)为二维连续型随机向量, 则(2.4.1)等价于

$$p(x,y)=p_\xi(x)p_\eta(y)\ \ a.e. \tag{2.4.3}$$

若(ξ,η)为二维连续型随机向量, 则(2.4.1)等价于

$$p(x,y)=p_\xi(x)p_\eta(y)\ \ a.e. \tag{2.4.3}$$

证明: 对任意的$x,y\in\mathbb{R}$,

$$\begin{aligned}
& F(x,y)=F_\xi(x)F_\eta(y) \\
\Longleftrightarrow\ & \int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv=\int_{-\infty}^{x}p_\xi(u)du\int_{-\infty}^{y}p_\eta(v)dv \\
\Longleftrightarrow\ & \int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv=\int_{-\infty}^{x}\int_{-\infty}^{y}p_\xi(u)p_\eta(v)dudv \\
\Longleftrightarrow\ & p(x,y)=p_\xi(x)p_\eta(y)\ \ a.e.
\end{aligned}$$

若(ξ,η)为二维连续型随机向量, 则(2.4.1)等价于

$$p(x,y)=p_\xi(x)p_\eta(y)\ \ a.e. \tag{2.4.3}$$

证明: 对任意的$x,y\in\mathbb{R}$,

$$\begin{aligned}
&F(x,y)=F_\xi(x)F_\eta(y)\\
\Longleftrightarrow\ &\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv=\int_{-\infty}^{x}p_\xi(u)du\int_{-\infty}^{y}p_\eta(v)dv\\
\Longleftrightarrow\ &\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv=\int_{-\infty}^{x}\int_{-\infty}^{y}p_\xi(u)p_\eta(v)dudv\\
\Longleftrightarrow\ &p(x,y)=p_\xi(x)p_\eta(y)\ \ a.e.
\end{aligned}$$

前面, 我们已经知道联合概率分布决定边际概率分布. 现在, 由(2.4.1), (2.4.2)和(2.4.3)可知: 当两个随机变量相互独立的时候, 边际概率分布和联合概率分布可以相互决定.

例2 设$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$, 求$\xi,\eta$独立的充要条件.

例2 设$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$，求$\xi,\eta$独立的充要条件.

解：因为$\xi\sim N(a,\sigma_1^2)$，$\eta\sim N(b,\sigma_2^2)$，所以

$$
\begin{aligned}
& \xi,\eta\text{独立} \\
\Longleftrightarrow\ & p(x,y)=p_\xi(x)p_\eta(y),a.e. \\
\Longleftrightarrow\ & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\cdot \\
& \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-a)^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\right]\right\} \\
& =\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-a)^2}{2\sigma_1^2}\right\}\cdot\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{(y-b)^2}{2\sigma_2^2}\right\},a.e. \\
\Longleftrightarrow\ &
\end{aligned}
$$

例2 设$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$, 求$\xi,\eta$独立的充要条件.

解: 因为$\xi\sim N(a,\sigma_1^2)$, $\eta\sim N(b,\sigma_2^2)$, 所以

$$
\begin{aligned}
& \xi,\eta\text{独立} \\
\Longleftrightarrow\quad & p(x,y)=p_\xi(x)p_\eta(y), a.e. \\
\Longleftrightarrow\quad & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\cdot \\
& \exp\Big\{-\frac{1}{2(1-r^2)}\Big[\frac{(x-a)^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\Big]\Big\} \\
& =\frac{1}{\sqrt{2\pi}\sigma_1}\exp\Big\{-\frac{(x-a)^2}{2\sigma_1^2}\Big\}\cdot\frac{1}{\sqrt{2\pi}\sigma_2}\exp\Big\{-\frac{(y-b)^2}{2\sigma_2^2}\Big\}, a.e. \\
\Longleftrightarrow\quad & r=0.
\end{aligned}
$$

n个随机变量的独立性:

n个随机变量的独立性:

定义

假设$F(x_1,\cdots,x_n)$, $F_1(x_1),\cdots,F_n(x_n)$分别为$\xi_1,\cdots,\xi_n$的联合分布函数和边际分布函数, 则当

$$F(x_1,\cdots,x_n)=F_1(x_1)\cdots F_n(x_n),\ \ \forall x_1,\cdots,x_n\in\mathbb{R}$$

时称$\xi_1,\cdots,\xi_n$相互独立.

推论

若$\xi_1, \cdots, \xi_n$相互独立, 则其中的任意$r(2 \le r \le n)$个也相互独立.

推论

若$\xi_1,\cdots,\xi_n$相互独立, 则其中的任意$r(2\le r\le n)$个也相互独立.

证明: 对任意的$x_{i_1},\cdots,x_{i_r}\in\mathbb{R}$,

$$\begin{aligned}
&\mathsf{P}(\xi_{i_1}\le x_{i_1},\cdots,\xi_{i_r}\le x_{i_r})\\
=&\mathsf{P}(\xi_{i_1}\le x_{i_1},\cdots,\xi_{i_r}\le x_{i_r},\xi_{i_{r+1}}<\infty,\cdots,\xi_{i_n}<\infty)\\
=&\mathsf{P}(\xi_{i_1}\le x_{i_1})\cdots\mathsf{P}(\xi_{i_r}\le x_{i_r})\times 1\times\cdots\times 1\\
=&\mathsf{P}(\xi_{i_1}\le x_{i_1})\cdots\mathsf{P}(\xi_{i_r}\le x_{i_r}).
\end{aligned}$$

得证.

无穷多个随机变量$\{\xi_i; i \geq 1\}$的独立性:

无穷多个随机变量$\{\xi_i; i \geq 1\}$的独立性:

定义

对于无穷多个随机变量$\{\xi_i; i \geq 1\}$, 若其中任意有限个随机变量都是相互独立的, 则称$\{\xi_i; i \geq 1\}$相互独立.

前面我们列举了随机变量独立性的各种表达形式, 有些是对一般随机变量成立的, 有些只对离散型或连续型才成立, 一般说来, 这些条件比较易于验证. 下面我们介绍另外一个条件, 这条件不易验证, 但在理论研究中有用.

前面我们列举了随机变量独立性的各种表达形式, 有些是对一般随机变量成立的, 有些只对离散型或连续型才成立, 一般说来, 这些条件比较易于验证. 下面我们介绍另外一个条件, 这条件不易验证, 但在理论研究中有用.

随机变量$\xi_1, \cdots, \xi_n$相互独立的充要条件是对一切一维Borel集$B_1, \cdots, B_n$, 下式成立:

$$\mathsf{P}(\xi_1 \in B_1, \cdots, \xi_n \in B_n) = \mathsf{P}(\xi_1 \in B_1) \cdots \mathsf{P}(\xi_n \in B_n).$$

前面我们列举了随机变量独立性的各种表达形式, 有些是对一般随机变量成立的, 有些只对离散型或连续型才成立, 一般说来, 这些条件比较易于验证. 下面我们介绍另外一个条件, 这条件不易验证, 但在理论研究中有用.

随机变量$\xi_1, \cdots, \xi_n$相互独立的充要条件是对一切一维Borel集$B_1, \cdots, B_n$, 下式成立:

$$\mathsf{P}(\xi_1 \in B_1, \cdots, \xi_n \in B_n) = \mathsf{P}(\xi_1 \in B_1) \cdots \mathsf{P}(\xi_n \in B_n).$$

n维随机向量$\boldsymbol{\xi}$与m维随机向量$\boldsymbol{\eta}$相互独立的充要条件是对任意的n维和m维Borel集A, B, 下式成立:

$$\mathsf{P}(\boldsymbol{\xi} \in A, \boldsymbol{\eta} \in B) = \mathsf{P}(\boldsymbol{\xi} \in A)\mathsf{P}(\boldsymbol{\eta} \in B).$$

例3 设ξ是只取常数a的退化分布. 求证: 对于任意随机变量η, ξ与η相互独立.

例3 设ξ是只取常数a的退化分布. 求证: 对于任意随机变量η, ξ与η相互独立.

证明: 易知ξ的分布函数为

$$F_\xi(x) = \begin{cases} 0, & x < a, \\ 1, & x \geq a. \end{cases}$$

例3 设ξ是只取常数a的退化分布. 求证: 对于任意随机变量η, ξ与η相互独立.

证明: 易知ξ的分布函数为

$$F_\xi(x) = \begin{cases} 0, & x < a, \\ 1, & x \geq a. \end{cases}$$

当$x < a$时,

$$\{\xi \leq x\} = \phi, \quad F_\xi(x) = 0.$$

所以, 对任意的$y \in \mathbb{R}$,

$$F(x, y) = \mathsf{P}(\xi \leq x, \eta \leq y) = 0 = F_\xi(x) F_\eta(y);$$

当$x \geq a$时,

$$\{\xi \leq x\} = \Omega, \quad F_\xi(x) = 1.$$

因此, 对任意的$y \in \mathbb{R}$,

$$F(x, y) = \mathsf{P}(\xi \leq x, \eta \leq y) = \mathsf{P}(\eta \leq y) = F_\eta(y) = F_\xi(x)F_\eta(y).$$

当$x \geq a$时,

$$\{\xi \leq x\} = \Omega, \quad F_\xi(x) = 1.$$

因此, 对任意的$y \in \mathbb{R}$,

$$F(x,y) = \mathsf{P}(\xi \leq x, \eta \leq y) = \mathsf{P}(\eta \leq y) = F_\eta(y) = F_\xi(x)F_\eta(y).$$

综上所述, 对任意的$x, y \in \mathbb{R}$, 均有

$$F(x,y) = F_\xi(x)F_\eta(y).$$

所以ξ与η相互独立.

条件分布

我们在第一章曾讨论过事件的条件概率，同样也可以考虑一个随机变量的条件分布(conditional distribution)，其条件与另一个随机变量的取值有关. 讨论从离散型随机变量开始.

条件分布

我们在第一章曾讨论过事件的条件概率, 同样也可以考虑一个随机变量的条件分布(conditional distribution), 其条件与另一个随机变量的取值有关. 讨论从离散型随机变量开始.

2.5.1 离散型的情形

条件分布

我们在第一章曾讨论过事件的条件概率, 同样也可以考虑一个随机变量的条件分布(conditional distribution), 其条件与另一个随机变量的取值有关. 讨论从离散型随机变量开始.

2.5.1 离散型的情形

设(ξ, η)是离散型随机向量. 分布列为

$$\mathrm{P}(\xi = x_i, \eta = y_j) = p_{ij}, \quad i, j = 1, 2, \cdots .$$

若已知$\xi = x_i$且$\mathrm{P}(\xi = x_i) > 0$, 则

$$\mathrm{P}(\eta = y_j | \xi = x_i) =$$

条件分布

我们在第一章曾讨论过事件的条件概率, 同样也可以考虑一个随机变量的条件分布(conditional distribution), 其条件与另一个随机变量的取值有关. 讨论从离散型随机变量开始.

2.5.1 离散型的情形

设(ξ, η)是离散型随机向量. 分布列为

$$\mathrm{P}(\xi = x_i, \eta = y_j) = p_{ij}, \quad i, j = 1, 2, \cdots .$$

若已知$\xi = x_i$且$\mathrm{P}(\xi = x_i) > 0$, 则

$$\mathrm{P}(\eta = y_j | \xi = x_i) = \frac{\mathrm{P}(\xi = x_i, \eta = y_j)}{\mathrm{P}(\xi = x_i)} = \frac{p_{ij}}{p_{i\cdot}}, \ j = 1, 2, \cdots \quad (2.5.1)$$

显然, (2.5.1)满足分布列的两个条件:

- $0 \le \dfrac{p_{ij}}{p_{i\cdot}} \le 1,\ j = 1, 2, \cdots;$
- $\displaystyle\sum_{j=1}^{\infty} \frac{p_{ij}}{p_{i\cdot}} = 1.$

显然, (2.5.1)满足分布列的两个条件:

- $0 \le \dfrac{p_{ij}}{p_{i\cdot}} \le 1,\ j = 1, 2, \cdots;$
- $\displaystyle\sum_{j=1}^{\infty} \frac{p_{ij}}{p_{i\cdot}} = 1.$

定义

称(2.5.1)为在$\xi = x_i$的条件下η的条件概率分布列, 简称为条件分布(列), 记作$p_{\eta|\xi}(y_j|x_i)$. 称

$$P(\eta \le y|\xi = x_i) = \sum_{j:y_j \le y} p_{\eta|\xi}(y_j|x_i)$$

为在$\xi = x_i$的条件下η的条件分布函数.

同理, 若已知$\eta = y_j$且$\mathsf{P}(\eta = y_j) > 0$, 那么在$\eta = y_j$的条件下$\xi$的条件分布列为

$$p_{\xi|\eta}(x_i|y_j) = \mathsf{P}(\xi = x_i|\eta = y_j) = \frac{\mathsf{P}(\xi = x_i, \eta = y_j)}{\mathsf{P}(\eta = y_j)} = \frac{p_{ij}}{p_{\cdot j}}, \quad i = 1, 2, \cdots .$$

称

$$\mathsf{P}(\xi \leq x|\eta = y_j) = \sum_{i: x_i \leq x} p_{\xi|\eta}(x_i|y_j)$$

为在$\eta = y_j$的条件下ξ的条件分布函数.

例1 在独立重复Bernoulli试验中，记p为每次试验“成功”的概率，S_n表示第n次成功时的试验次数．求
(1) 在$S_n = t$的条件下S_{n+1}的条件概率分布;
(2) 在$S_{n+1} = w$的条件下S_n的条件概率分布．

例1 在独立重复Bernoulli试验中, 记p为每次试验“成功”的概率, S_n表示第n次成功时的试验次数. 求
(1) 在$S_n = t$的条件下S_{n+1}的条件概率分布;
(2) 在$S_{n+1} = w$的条件下S_n的条件概率分布.

解: 对$t < w$, 事件$\{S_n = t, S_{n+1} = w\}$意味着在w次试验中, 第t, w次试验出现“成功”, 在第1次到第$t-1$次试验中出现$n-1$次“成功”, 其余均出现“失败”.

例1 在独立重复Bernoulli试验中，记p为每次试验“成功”的概率，S_n表示第n次成功时的试验次数．求
(1) 在$S_n = t$的条件下S_{n+1}的条件概率分布;
(2) 在$S_{n+1} = w$的条件下S_n的条件概率分布.

解: 对$t < w$，事件$\{S_n = t, S_{n+1} = w\}$意味着在w次试验中，第t，w次试验出现“成功”，在第1次到第$t-1$次试验中出现$n-1$次“成功”，其余均出现“失败”. 所以

$$\begin{aligned} \mathrm{P}(S_n = t, S_{n+1} = w) &= p \cdot p \cdot \binom{t-1}{n-1} p^{n-1} q^{w-(n+1)} \\ &= \binom{t-1}{n-1} p^{n+1} q^{w-(n+1)}. \end{aligned}$$

从而在$S_n = t$的条件下S_{n+1}的条件概率分布为

$$\mathsf{P}(S_{n+1} = w | S_n = t) = \frac{\mathsf{P}(S_n = t, S_{n+1} = w)}{\mathsf{P}(S_n = t)} = pq^{w-t-1}, \ w = t+1, \cdots .$$

从而在$S_n = t$的条件下S_{n+1}的条件概率分布为

$$\mathsf{P}(S_{n+1} = w | S_n = t) = \frac{\mathsf{P}(S_n = t, S_{n+1} = w)}{\mathsf{P}(S_n = t)} = pq^{w-t-1},\ w = t+1, \cdots .$$

又

$$\mathsf{P}(S_{n+1} = w | S_n = t) = \mathsf{P}(S_{n+1} - S_n = w - t | S_n = t),$$

这意味着在$S_n = t$的条件下$S_{n+1} - S_n$服从参数为p的几何分布.

在$S_{n+1}=w$的条件下S_n的条件概率分布列为

$$\begin{aligned}
\mathrm{P}(S_n=t|S_{n+1}=w)=&\frac{\mathrm{P}(S_n=t,S_{n+1}=w)}{\mathrm{P}(S_{n+1}=w)}\\
=&\frac{\binom{t-1}{n-1}p^{n+1}q^{w-(n+1)}}{\binom{w-1}{n}p^{n+1}q^{w-(n+1)}}\\
=&\frac{\binom{t-1}{n-1}}{\binom{w-1}{n}},\quad t=n,\cdots,w-1.
\end{aligned}$$

这一条件概率分布不依赖于p.

在$S_{n+1}=w$的条件下S_n的条件概率分布列为

$$\begin{aligned}\mathrm{P}(S_n=t|S_{n+1}=w)=&\frac{\mathrm{P}(S_n=t,S_{n+1}=w)}{\mathrm{P}(S_{n+1}=w)}\\=&\frac{\binom{t-1}{n-1}p^{n+1}q^{w-(n+1)}}{\binom{w-1}{n}p^{n+1}q^{w-(n+1)}}\\=&\frac{\binom{t-1}{n-1}}{\binom{w-1}{n}},\quad t=n,\cdots,w-1.\end{aligned}$$

这一条件概率分布不依赖于p.

2.5.2 连续型的情形

2.5.2 连续型的情形

设(ξ, η)有联合密度函数$p(x, y)$和联合分布函数$F(x, y)$. 因为对任何x, $\mathsf{P}(\xi = x) = 0$, 条件概率$\mathsf{P}(\eta \leq y | \xi = x)$没有定义, 故只能借助于密度函数.

若$p_\xi(x)>0$, 定义

$$
\begin{aligned}
\mathsf{P}(\eta\le y|\xi=x) &= \lim_{\Delta x\to 0}\mathsf{P}(\eta\le y|x<\xi\le x+\Delta x)\\
&= \lim_{\Delta x\to 0}\frac{\mathsf{P}(x<\xi\le x+\Delta x,\eta\le y)}{\mathsf{P}(x<\xi\le x+\Delta x)}\\
&= \lim_{\Delta x\to 0}\frac{F(x+\Delta x,y)-F(x,y)}{F_\xi(x+\Delta x)-F_\xi(x)}\\
&= \lim_{\Delta x\to 0}\frac{\frac{F(x+\Delta x,y)-F(x,y)}{\Delta x}}{\frac{F_\xi(x+\Delta x)-F_\xi(x)}{\Delta x}}\\
&= \frac{\frac{\partial F(x,y)}{\partial x}}{F'_\xi(x)} = \frac{\int_{-\infty}^{y}p(x,v)\mathrm{d}v}{p_\xi(x)}\\
&= \int_{-\infty}^{y}\frac{p(x,v)}{p_\xi(x)}\mathrm{d}v.
\end{aligned}
$$

若$p_\xi(x) > 0$, 定义

$$
\begin{aligned}
\mathsf{P}(\eta \le y | \xi = x) &= \lim_{\Delta x \to 0} \mathsf{P}(\eta \le y | x < \xi \le x + \Delta x) \\
&= \lim_{\Delta x \to 0} \frac{\mathsf{P}(x < \xi \le x + \Delta x, \eta \le y)}{\mathsf{P}(x < \xi \le x + \Delta x)} \\
&= \lim_{\Delta x \to 0} \frac{F(x + \Delta x, y) - F(x, y)}{F_\xi(x + \Delta x) - F_\xi(x)} \\
&= \lim_{\Delta x \to 0} \frac{\frac{F(x+\Delta x, y) - F(x,y)}{\Delta x}}{\frac{F_\xi(x+\Delta x) - F_\xi(x)}{\Delta x}} \\
&= \frac{\frac{\partial F(x,y)}{\partial x}}{F_\xi'(x)} = \frac{\int_{-\infty}^{y} p(x, v)\mathrm{d}v}{p_\xi(x)} \\
&= \int_{-\infty}^{y} \frac{p(x, v)}{p_\xi(x)} \mathrm{d}v.
\end{aligned}
$$

显然，$\frac{p(x,v)}{p_\xi(x)}$满足概率密度函数的两个条件:

- $\dfrac{p(x,v)}{p_\xi(x)} \geq 0$;
- $\displaystyle\int_{-\infty}^{\infty} \frac{p(x,v)}{p_\xi(x)} \mathrm{d}v = 1$.

定义

设(ξ,η)有联合密度函数$p(x,y)$, ξ有边际密度函数$p_\xi(x)=\int_{-\infty}^{\infty}p(x,y)\mathrm{d}y$. 若在$x$处, $p_\xi(x)>0$, 则称

$$P(\eta\le y|\xi=x)=\int_{-\infty}^{y}\frac{p(x,v)}{p_\xi(x)}\mathrm{d}v,\quad y\in\mathbb{R}$$

为在$\xi=x$的条件下, η的条件分布函数, 简称为条件分布, 记作$F_{\eta|\xi}(y|x)$. 称

$$p_{\eta|\xi}(y|x)=\frac{p(x,y)}{p_\xi(x)},\quad y\in\mathbb{R} \tag{2.5.2}$$

为在$\xi=x$的条件下, η的条件概率密度函数, 简称为条件密度函数.

定义

设(ξ,η)有联合密度函数$p(x,y)$, ξ有边际密度函数$p_\xi(x)=\int_{-\infty}^{\infty}p(x,y)\mathrm{d}y$. 若在$x$处, $p_\xi(x)>0$, 则称

$$P(\eta\le y|\xi=x)=\int_{-\infty}^{y}\frac{p(x,v)}{p_\xi(x)}\mathrm{d}v,\quad y\in\mathbb{R}$$

为在$\xi=x$的条件下, η的条件分布函数, 简称为条件分布, 记作$F_{\eta|\xi}(y|x)$. 称

$$p_{\eta|\xi}(y|x)=\frac{p(x,y)}{p_\xi(x)},\quad y\in\mathbb{R} \tag{2.5.2}$$

为在$\xi=x$的条件下, η的条件概率密度函数, 简称为条件密度函数.

同理, 若$p_\eta(y) > 0$, 定义$\eta = y$时ξ的条件概率密度函数为

$$p_{\xi|\eta}(x|y) = \frac{p(x,y)}{p_\eta(y)}, \quad x \in \mathbb{R}.$$

同理, 若$p_\eta(y) > 0$, 定义$\eta = y$时ξ的条件概率密度函数为

$$p_{\xi|\eta}(x|y) = \frac{p(x,y)}{p_\eta(y)}, \quad x \in \mathbb{R}.$$

由条件概率密度公式可得

$$p(x,y) = p_{\xi|\eta}(x|y)p_\eta(y) = p_{\eta|\xi}(y|x)p_\xi(x),$$

从而有

$$p_{\eta|\xi}(y|x) = \frac{p_{\xi|\eta}(x|y)p_\eta(y)}{\int_{-\infty}^{\infty} p_{\xi|\eta}(x|v)p_\eta(v)\mathrm{d}v}.$$

这可看做是贝叶斯公式的连续型形式.

例2 设$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$, 求条件密度函数$p_{\eta|\xi}(y|x)$.

例2 设$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$, 求条件密度函数$p_{\eta|\xi}(y|x)$.

解: 我们已经知道: (ξ,η)的联合密度函数为

$$
\begin{aligned}
& p(x,y) \\
=& \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\cdot \\
& \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-a)^2}{\sigma_1^2}-\frac{2r(x-a)(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\right]\right\} \\
=& \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-a)^2}{2\sigma_1^2}\right\}. \\
& \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{[y-b-\frac{r\sigma_2}{\sigma_1}(x-a)]^2}{2\sigma_2^2(1-r^2)}\right\}, \\
& \qquad -\infty<x<\infty,\ -\infty<y<\infty;
\end{aligned}
$$

ξ的边际密度函数为

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{ -\frac{(x-a)^2}{2\sigma_1^2} \right\}, \quad -\infty < x < \infty.$$

ξ的边际密度函数为

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{ -\frac{(x-a)^2}{2\sigma_1^2} \right\}, \quad -\infty < x < \infty.$$

所以条件概率密度函数$p_{\eta|\xi}(y|x)$为

$$\begin{aligned} p_{\eta|\xi}(y|x) =& \frac{p(x,y)}{p_\xi(x)} \\ =& \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} \exp\left\{ -\frac{[y-b-\frac{r\sigma_2}{\sigma_1}(x-a)]^2}{2\sigma_2^2(1-r^2)} \right\}, \\ & -\infty < y < \infty. \end{aligned}$$

即

$$\eta|_{\xi=x} \sim N(b + \frac{r\sigma_2}{\sigma_1}(x-a), (1-r^2)\sigma_2^2).$$

2.5.3 一般情形

2.5.3 一般情形

一般地, 设(ξ,η)为二维随机向量, 对给定的x, 如果极限

$$\lim_{\varepsilon\to 0^+}\frac{\mathsf{P}(\eta\le y, x<\xi\le x+\varepsilon)}{\mathsf{P}(x<\xi\le x+\varepsilon)}$$

对任意的$y\in\mathbb{R}$均存在, 则称此极限

$$F_{\eta|\xi}(y|x):=\lim_{\varepsilon\to 0^+}\frac{\mathsf{P}(\eta\le y, x<\xi\le x+\varepsilon)}{\mathsf{P}(x<\xi\le x+\varepsilon)}$$

为在$\xi=x$的条件下η的条件分布函数, 简称为条件分布.

如果存在$y_j, j=1,2,\cdots$, 使得$F_{\eta|\xi}(y|x)$能表示为

$$F_{\eta|\xi}(y|x)=\sum_{j:y_j\leq y} p_{\eta|\xi}(y_j|x), \quad y\in\mathbb{R},$$

则称$p_{\eta|\xi}(y_j|x), j=1,2,\cdots$为条件分布列.

如果存在$y_j, j=1,2,\cdots$，使得$F_{\eta|\xi}(y|x)$能表示为

$$F_{\eta|\xi}(y|x) = \sum_{j:y_j\le y} p_{\eta|\xi}(y_j|x),\quad y\in\mathbb{R},$$

则称$p_{\eta|\xi}(y_j|x), j=1,2,\cdots$为条件分布列.

如果$F_{\eta|\xi}(y|x)$能表示为

$$F_{\eta|\xi}(y|x) = \int_{-\infty}^{y} p_{\eta|\xi}(v|x)\mathrm{d}v,\quad y\in\mathbb{R},$$

则称$p_{\eta|\xi}(y|x)$为条件概率密度函数.

*例3 设Λ服从Γ分布$\Gamma(b,a)$, 在条件$\Lambda=\lambda$ 下, X服从参数为λ的Poisson分布. 求在$X=x$条件下Λ的分布.

*例3 设Λ服从Γ分布$\Gamma(b,a)$, 在条件$\Lambda=\lambda$ 下, X服从参数为λ的Poisson分布. 求在$X=x$条件下Λ的分布.

解: Λ为连续型随机变量, X为离散型随机变量. 对$x=0,1,\cdots$, 有

$$\mathsf{P}(X=x|\Lambda=\lambda)=\frac{\lambda^x}{x!}e^{-\lambda}.$$

由定义

$$\mathsf{P}(X=x|\Lambda=\lambda)=\lim_{\Delta\lambda\to 0}\frac{\mathsf{P}(X=x,\Lambda\in(\lambda,\lambda+\Delta\lambda])}{\mathsf{P}(\Lambda\in(\lambda,\lambda+\Delta\lambda])}.$$

即

$$\begin{aligned}&\mathsf{P}(X=x,\Lambda\in(\lambda,\lambda+\Delta\lambda])\\=&\mathsf{P}(X=x|\Lambda=\lambda)\mathsf{P}(\Lambda\in(\lambda,\lambda+\Delta\lambda])+o(\Delta\lambda)\\=&\mathsf{P}(X=x|\Lambda=\lambda)p_\Lambda(\lambda)\Delta\lambda+o(\Delta\lambda).\end{aligned}$$

所以

$$\mathsf{P}(X=x,\Lambda\le y)=\int_{-\infty}^{y}\mathsf{P}(X=x|\Lambda=\lambda)p_\Lambda(\lambda)\mathrm{d}\lambda.$$

从而

$$\mathsf{P}(\Lambda \le y | X = x) = \frac{\mathsf{P}(X = x, \Lambda \le y)}{\mathsf{P}(X = x)} = \int_{-\infty}^{y} \frac{\mathsf{P}(X = x | \Lambda = \lambda) p_\Lambda(\lambda)}{\mathsf{P}(X = x)} \mathrm{d}\lambda.$$

因此, 在$X = x$的条件下, Λ的概率密度函数为

$$\begin{aligned} p_{\Lambda|X}(\lambda|x) &= \frac{\mathsf{P}(X = x | \Lambda = \lambda) p_\Lambda(\lambda)}{\mathsf{P}(X = x)} \\ &\propto_\lambda \lambda^x e^{-\lambda} \lambda^{b-1} e^{-\lambda a} \\ &\propto_\lambda \lambda^{x+b-1} e^{-(a+1)\lambda}, \quad \lambda > 0. \end{aligned}$$

从而

$$\mathsf{P}(\Lambda \le y | X = x) = \frac{\mathsf{P}(X = x, \Lambda \le y)}{\mathsf{P}(X = x)} = \int_{-\infty}^{y} \frac{\mathsf{P}(X = x | \Lambda = \lambda) p_{\Lambda}(\lambda)}{\mathsf{P}(X = x)} \mathrm{d}\lambda.$$

因此, 在$X = x$的条件下, Λ的概率密度函数为

$$\begin{aligned} p_{\Lambda|X}(\lambda|x) &= \frac{\mathsf{P}(X = x|\Lambda = \lambda)p_{\Lambda}(\lambda)}{\mathsf{P}(X = x)} \\ &\propto_{\lambda} \lambda^{x} e^{-\lambda} \lambda^{b-1} e^{-\lambda a} \\ &\propto_{\lambda} \lambda^{x+b-1} e^{-(a+1)\lambda}, \quad \lambda > 0. \end{aligned}$$

将上式右边添加正则化常数因子使其积分为1, 得

$$p_{\Lambda|X}(\lambda|x) = \frac{(a+1)^{x+b}}{\Gamma(x+b)} \lambda^{x+b-1} e^{-(a+1)\lambda}, \quad \lambda > 0.$$

即在$X = x$条件下, Λ服从Γ分布$\Gamma(a+1, x+b)$.

*2.5.4 给定随机变量下的条件概率

*2.5.4 给定随机变量下的条件概率

设X为随机变量，D是其值域，A为随机事件，则

$$g(x)=\lim_{\varepsilon\to 0^+}\frac{\mathsf{P}(A, x-\varepsilon<X\leq x+\varepsilon)}{\mathsf{P}(x-\varepsilon<X\leq x+\varepsilon)},\quad x\in D$$

是定义在D上的实函数，记作$\mathsf{P}(A|X=x)$. 于是可以定义随机变量$g(X)$，我们称$g(X)$为事件A关于随机变量X的条件概率，记作$\mathsf{P}(A|X)$.

值得指出，作为X的函数，条件概率$\mathsf{P}(A|X)$ 是随机变量.

根据上述定义

$$\mathsf{P}(A|X) = g(X) \text{ 当且仅当 } \mathsf{P}(A|X = x) = g(x),\ x \in D.$$

根据上述定义

$$\mathsf{P}(A|X) = g(X) \text{ 当且仅当 } \mathsf{P}(A|X = x) = g(x),\ x \in D.$$

如果X为离散型随机变量，其概率分布为$p_X(x_i)$，那么由全概率公式有

$$\begin{aligned}\mathsf{P}(A) &= \sum_i \mathsf{P}(A|X = x_i)\mathsf{P}(X = x_i)\\ &= \sum_i \mathsf{P}(A|X = x_i)p_X(x_i)\\ &= g(x_i)p_X(x_i).\end{aligned}$$

当X为连续型随机变量, 记其密度函数为$p_X(x)$. 这时由于

$$\mathsf{P}(A|X=x) \approx \frac{\mathsf{P}(A, X \in (x, x+\Delta x])}{\mathsf{P}(X \in (x, x+\Delta x])}.$$

即

$$\mathsf{P}(A, X \in (x, x+\Delta x]) = \mathsf{P}(A|X=x)p_X(x)\Delta x + o(\Delta x).$$

因此

$$\begin{aligned}\mathsf{P}(A) =& \mathsf{P}(A, -\infty < X < \infty) \\ =& \int_{-\infty}^{\infty} \mathsf{P}(A|X=x)p_X(x)\mathrm{d}x \\ =& \int_{-\infty}^{\infty} g(x)p_X(x)\mathrm{d}x.\end{aligned}$$

这就是全概率公式的连续形式.

例4 设$U_1, U_2, \cdots$为一列独立随机变量, 均服从$(0,1)$上的均匀分布. 令

$$\xi = \min\{n \geq 1 : U_1 + \cdots + U_n > 1\}.$$

求ξ的概率分布.

例4 设$U_1, U_2, \cdots$为一列独立随机变量, 均服从$(0,1)$上的均匀分布. 令

$$\xi = \min\{n \geq 1 : U_1 + \cdots + U_n > 1\}.$$

求ξ的概率分布.

解: 我们来求概率$\mathsf{P}(\xi > n)$. 考虑更一般的情形, 对$0 < x < 1$, 记

$$\xi(x) = \min\{n \geq 1 : U_1 + \cdots + U_n > x\},$$

则容易看出

$$\mathsf{P}(\xi(x) > 1) = \mathsf{P}(U_1 \leq x) = x.$$

进一步, 我们有下列的递推关系:

$$\begin{aligned}
\mathsf{P}(\xi(x) > n+1) =& \mathsf{P}(U_1 + \cdots + U_{n+1} \le x) \\
=& \int_{-\infty}^{\infty} \mathsf{P}(U_1 + \cdots + U_{n+1} \le x | U_1 = y) p_{U_1}(y) \mathrm{d}y \\
=& \int_0^1 \mathsf{P}(y + U_2 + \cdots + U_{n+1} \le x | U_1 = y) \mathrm{d}y \\
=& \int_0^x \mathsf{P}(U_1 + \cdots + U_n \le x - y) \mathrm{d}y \\
=& \int_0^x \mathsf{P}(U_1 + \cdots + U_n \le u) \mathrm{d}u \\
=& \int_0^x \mathsf{P}(\xi(x) > n) \mathrm{d}u.
\end{aligned}$$

由归纳法得
$$\mathsf{P}(\xi(x) > n) = \frac{x^n}{n!}.$$
特别地，当$x = 1$时，$\mathsf{P}(\xi > n) = 1/n!$. 这样就我们得到$\xi$的分布列为
$$\mathsf{P}(\xi = n) = \frac{1}{(n-1)!} - \frac{1}{n!}, \quad n = 1, 2, \cdots.$$

随机变量的函数及其分布

在许多问题中需要计算随机变量的函数的概率分布.

随机变量的函数及其分布

在许多问题中需要计算随机变量的函数的概率分布.

例如在无线电接收中, 某时刻收到的信号是一个随机变量ξ, 若我们把这个信号通过平方检波器, 则输出的信号为$\eta = \xi^2$, 这时我们需要求η的概率分布.

随机变量的函数及其分布

在许多问题中需要计算随机变量的函数的概率分布.

例如在无线电接收中, 某时刻收到的信号是一个随机变量ξ, 若我们把这个信号通过平方检波器, 则输出的信号为$\eta = \xi^2$, 这时我们需要求η的概率分布.

又如在统计物理中, 已知分子运动速度ξ的分布, 要求其动能$\eta = \frac{1}{2}m\xi^2$的概率分布.

假设ξ是一个随机变量, $y = g(x)$是一个实函数, 那么$\eta = g(\xi)$是ξ的函数.

假设ξ是一个随机变量，$y = g(x)$是一个实函数，那么$\eta = g(\xi)$是ξ的函数.

问题:

- $\eta = g(\xi)$是随机变量吗?
- 如果η是随机变量，那么η的概率分布与ξ的概率分布有何关系?

对于$\eta = g(\xi)$, Borel集$B \in \mathscr{B}$,

$$\begin{aligned}\{\omega : \eta(\omega) \in B\} =& \{\omega : g(\xi(\omega)) \in B\} \\ =& \Big\{\omega : \xi(\omega) \in \{x : g(x) \in B\}\Big\}.\end{aligned}$$

为了使得η是一个随机变量, 我们要求对任意的Borel集$B \in \mathscr{B}$,

$$\Big\{\omega : \xi(\omega) \in \{x : g(x) \in B\}\Big\}$$

是一个来自事件域的事件. 因此, 只需对任意的Borel集$B \in \mathscr{B}$, 使得

$$\{x : g(x) \in B\}$$

也是一个Borel集.

定义

设$g(x)$是一维实函数, $\mathscr{B}$是$\mathbb{R}$上的Borel域. 若对任意的$B \in \mathscr{B}$, 都有

$$\{x : g(x) \in B\} = g^{-1}(B) \in \mathscr{B}$$

(即当$g(x)$的值域是Borel集时, 其原像也是Borel集). 则称$g(x)$是一元Borel函数.

定义

设$g(x)$是一维实函数, $\mathscr{B}$是$\mathbb{R}$上的Borel域. 若对任意的$B \in \mathscr{B}$, 都有

$$\{x : g(x) \in B\} = g^{-1}(B) \in \mathscr{B}$$

(即当$g(x)$的值域是Borel集时, 其原像也是Borel集). 则称$g(x)$是一元Borel函数.

注: 实变函数论中可以证明: 一切分段连续、分段单调的函数都是Borel函数, 故它是一类十分广泛的函数.

由以上分析可知，若ξ是定义在概率空间$(\Omega, \mathscr{F}, \mathsf{P})$上的随机变量，函数$f(x)$是一元Borel函数，则$\eta = f(\xi)$是随机变量.

由以上分析可知, 若ξ是定义在概率空间$(\Omega, \mathscr{F}, \mathsf{P})$上的随机变量, 函数$f(x)$是一元Borel函数, 则$\eta = f(\xi)$是随机变量.

类似地, 我们可以定义n元Borel函数(略). 若$f(x_1, \cdots, x_n)$是n元Borel函数, 则$\eta = f(\xi_1, \cdots, \xi_n)$就是随机变量.

由以上分析可知，若ξ是定义在概率空间$(\Omega, \mathscr{F}, \mathrm{P})$上的随机变量，函数$f(x)$是一元Borel函数，则$\eta = f(\xi)$是随机变量.

类似地，我们可以定义n元Borel函数(略). 若$f(x_1, \cdots, x_n)$是n元Borel函数，则$\eta = f(\xi_1, \cdots, \xi_n)$就是随机变量.

不作其它说明的话，今后我们讨论的随机变量的函数都是指一元或多元的Borel函数.

2.6.1 离散型随机变量的函数

2.6.1 离散型随机变量的函数

例1 设ξ的分布列为

ξ	-1	0	1	2
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$

,

$\eta = 2\xi - 1, \zeta = \xi^2$. 求$\eta, \zeta$各自的概率分布.

2.6.1 离散型随机变量的函数

例1 设ξ的分布列为

ξ	-1	0	1	2
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$

,

$\eta = 2\xi - 1, \zeta = \xi^2$. 求$\eta, \zeta$各自的概率分布.

解: η的所有可能取值为$-3, -1, 1, 3$. 它们相应的概率为

$$
\begin{aligned}
\mathrm{P}(\eta = -3) = \mathrm{P}(\xi = -1) = \frac{1}{4}, \\
\mathrm{P}(\eta = -1) = \mathrm{P}(\xi = 0) = \frac{1}{2}, \\
\mathrm{P}(\eta = 1) = \mathrm{P}(\xi = 1) = \frac{1}{8}, \\
\mathrm{P}(\eta = 3) = \mathrm{P}(\xi = 2) = \frac{1}{8}.
\end{aligned}
$$

所以, η的分布列为

$$\begin{array}{c|cccc} \eta & -3 & -1 & 1 & 3 \\ \hline \mathsf{P} & \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{array}.$$

ζ的所有可能取值为: $0, 1, 4$. 它们相应的概率为

$$
\begin{aligned}
\mathsf{P}(\zeta = 0) &= \mathsf{P}(\xi = 0) = \frac{1}{2}, \\
\mathsf{P}(\zeta = 1) &= \mathsf{P}(\xi = -1) + \mathsf{P}(\xi = 1) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8}, \\
\mathsf{P}(\zeta = 4) &= \mathsf{P}(\xi = 2) = \frac{1}{8}.
\end{aligned}
$$

ζ的所有可能取值为: $0, 1, 4$. 它们相应的概率为

$$\mathrm{P}(\zeta = 0) = \mathrm{P}(\xi = 0) = \frac{1}{2},$$
$$\mathrm{P}(\zeta = 1) = \mathrm{P}(\xi = -1) + \mathrm{P}(\xi = 1) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8},$$
$$\mathrm{P}(\zeta = 4) = \mathrm{P}(\xi = 2) = \frac{1}{8}.$$

所以, ζ的分布列为

$$\begin{array}{c|ccc} \zeta & 0 & 1 & 4 \\ \hline \mathrm{P} & \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \end{array}.$$

例2 设$\xi \sim B(n_1, p), \eta \sim B(n_2, p)$, ξ, η相互独立. 求$\zeta = \xi + \eta$的概率分布.

例2 设$\xi \sim B(n_1,p), \eta \sim B(n_2,p)$, ξ,η相互独立. 求$\zeta=\xi+\eta$的概率分布.

解: 易知ζ的所有可能取值为: $0,1,\cdots,n_1+n_2$. 对应的概率为

$$
\begin{aligned}
\mathsf{P}(\zeta=r) &= \sum_{k=0}^{r} \mathsf{P}(\xi=k,\eta=r-k) \\
&= \sum_{k=0}^{r} \mathsf{P}(\xi=k)\mathsf{P}(\eta=r-k) \\
&= \sum_{k=0}^{r} \binom{n_1}{k} p^k q^{n_1-k} \binom{n_2}{r-k} p^{r-k} q^{n_2-r+k} \\
&= p^r q^{n_1+n_2-r} \sum_{k=0}^{r} \binom{n_1}{k}\binom{n_2}{r-k} \\
&= \binom{n_1+n_2}{r} p^r q^{n_1+n_2-r}, \quad r=0,1,\cdots,n_1+n_2.
\end{aligned}
$$

所以, $\zeta = \xi + \eta \sim B(n_1 + n_2, p)$. (二项分布的再生性/可加性(additive property))

所以, $\zeta = \xi + \eta \sim B(n_1 + n_2, p)$. (二项分布的再生性/可加性(additive property))

离散卷积公式:

$$\mathsf{P}(\xi + \eta = r) = \sum_{k=0}^{r} \mathsf{P}(\xi = k)\mathsf{P}(\eta = r - k).$$

2.6.2 一维连续型随机变量的函数的分布

2.6.2 一维连续型随机变量的函数的分布

设ξ的密度函数为$p(x)$. $\eta = f(\xi)$, $G(y)$是η的分布函数, 即

$$G(y) = \mathsf{P}(\eta \le y) = \mathsf{P}(f(\xi) \le y).$$

注意到

$$D := \{x : f(x) \le y\}$$

是一维Borel集(因为$f(x)$是一元Borel函数), 所以

$$G(y) = \mathsf{P}(\xi \in D) =$$

2.6.2 一维连续型随机变量的函数的分布

设ξ的密度函数为$p(x)$. $\eta = f(\xi)$, $G(y)$是η的分布函数, 即

$$G(y) = \mathsf{P}(\eta \le y) = \mathsf{P}(f(\xi) \le y).$$

注意到

$$D := \{x : f(x) \le y\}$$

是一维Borel集(因为$f(x)$是一元Borel函数), 所以

$$G(y) = \mathsf{P}(\xi \in D) = \int_D p(x)\mathrm{d}x.$$

注: 在一般情形下, 我们无法确定η也是连续型随机变量. 但在一些特殊的场合, 我们不但可以确定η是连续型随机变量, 而且还可以直接导出η的密度函数.

定理

若$f(x)$严格单调, 其反函数$f^{-1}(y)$有连续导函数. 设ξ的密度函数为$p(x)$. 则$\eta = f(\xi)$也是连续型随机变量, 其密度函数为

$$g(y) = \begin{cases} p(f^{-1}(y)) \cdot |\left(f^{-1}(y)\right)'|, & y \in f(x)\text{的值域}, \\ 0, & \text{其他}. \end{cases}$$

定理

若$f(x)$严格单调，其反函数$f^{-1}(y)$有连续导函数．设ξ的密度函数为$p(x)$．则$\eta=f(\xi)$也是连续型随机变量，其密度函数为

$$g(y)=\begin{cases} p(f^{-1}(y))\cdot\left|\left(f^{-1}(y)\right)'\right|, & y\in f(x)\text{的值域}, \\ 0, & \text{其他}. \end{cases}$$

证明：不妨设$f(x)$严格单调增加，$-\infty<x<\infty$时$A<f(x)<B$.

定理

若$f(x)$严格单调, 其反函数$f^{-1}(y)$有连续导函数. 设ξ的密度函数为$p(x)$. 则$\eta = f(\xi)$也是连续型随机变量, 其密度函数为

$$g(y) = \begin{cases} p(f^{-1}(y)) \cdot \left|\left(f^{-1}(y)\right)'\right|, & y \in f(x)\text{的值域}, \\ 0, & \text{其他}. \end{cases}$$

证明: 不妨设$f(x)$严格单调增加, $-\infty < x < \infty$时$A < f(x) < B$.
若$y \le A$, 则$G(y) = 0$, $p_\eta(y) = 0 = g(y)$;

定理

若$f(x)$严格单调, 其反函数$f^{-1}(y)$有连续导函数. 设ξ的密度函数为$p(x)$. 则$\eta = f(\xi)$也是连续型随机变量, 其密度函数为

$$g(y) = \begin{cases} p(f^{-1}(y)) \cdot |(f^{-1}(y))'|, & y \in f(x)\text{的值域}, \\ 0, & \text{其他}. \end{cases}$$

证明: 不妨设$f(x)$严格单调增加, $-\infty < x < \infty$时$A < f(x) < B$.
若$y \leq A$, 则$G(y) = 0$, $p_\eta(y) = 0 = g(y)$;
若$y \geq B$, 则$G(y) = 1$, $p_\eta(y) = 0 = g(y)$;

若$A < y < B$, 则

$$\{\eta \le y\} = \{f(\xi) \le y\} = \{\xi \le f^{-1}(y)\}.$$

因此,

$$G(y) = \mathsf{P}(\eta \le y) = \int_{-\infty}^{f^{-1}(y)} p(x)\mathrm{d}x.$$

令$x = f^{-1}(v)$, 得

$$G(y) = \int_A^y p(f^{-1}(v)) \cdot \left(f^{-1}(v)\right)' dv = \int_{-\infty}^y g(v)\mathrm{d}v.$$

综上所述, 可知当$f(x)$严格单调增加时, 结论成立.

若$A < y < B$, 则

$$\{\eta \le y\} = \{f(\xi) \le y\} = \{\xi \le f^{-1}(y)\}.$$

因此,

$$G(y) = \mathsf{P}(\eta \le y) = \int_{-\infty}^{f^{-1}(y)} p(x)\mathrm{d}x.$$

令$x = f^{-1}(v)$, 得

$$G(y) = \int_A^y p(f^{-1}(v)) \cdot \left(f^{-1}(v)\right)' dv = \int_{-\infty}^y g(v)\mathrm{d}v.$$

综上所述, 可知当$f(x)$严格单调增加时, 结论成立.

当$f(x)$严格单调递减时, 类似可证.

推论

若$y=f(x)$在不相重叠的区间$I_1, I_2, \cdots$上逐段严格单调, 在各段的反函数$h_1(y), h_2(y), \cdots$ 有连续导函数, 则$\eta=f(\xi)$是连续型随机变量, 其密度函数为

$$g(y)=\begin{cases} \sum_i p(h_i(y))\cdot|h_i'(y)|, & y\in \text{各}h_i(y)\text{的定义域}, \\ 0, & \text{其他}. \end{cases}$$

证明: 注意到

$$\{f(\xi) \leq y\} = \{\xi \in \sum_i E_i(y)\},$$

其中$E_i(y)$是I_i中满足$f(x) \leq y$的x的集合.

证明: 注意到

$$\{f(\xi)\leq y\}=\{\xi\in\sum_i E_i(y)\},$$

其中$E_i(y)$是I_i中满足$f(x)\leq y$的x的集合. 再仿照前面定理的证明, 可得当$y\in$各$h_i(y)$的定义域时,

$$\begin{aligned}\mathsf{P}(\eta\leq y)=&\mathsf{P}(\xi\in\sum_i E_i(y))=\sum_i\int_{E_i(y)}p(x)\mathrm{d}x\\ &=\sum_i\int_{A_i}^{y}p(h_i(x))|h_i^{'}(x)|\mathrm{d}x\\ &=\int_{-\infty}^{y}g(x)\mathrm{d}x.\end{aligned}$$

例3 设$\xi \sim N(a, \sigma^2)$, 求$\eta = k\xi + b$的密度函数$(k \neq 0)$.

例3 设$\xi \sim N(a, \sigma^2)$, 求$\eta = k\xi + b$的密度函数$(k \neq 0)$.

解: $y = kx + b$是严格单调的, 且$x = f^{-1}(y) = (y-b)/k$有连续导函数. 又ξ的密度函数为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty.$$

所以利用前面的定理可知, η的密度函数为

$$\begin{aligned} g(y) =& \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(\frac{y-b}{k}-a)^2}{2\sigma^2}\right\} \cdot \left|\frac{1}{k}\right| \\ =& \frac{1}{\sqrt{2\pi}|k|\sigma} \exp\left\{-\frac{(y-ka-b)^2}{2k^2\sigma^2}\right\}, \quad -\infty < y < \infty. \end{aligned}$$

即, $\eta \sim N(ka+b, k^2\sigma^2)$.

注: 若$\eta = (\xi - a)/\sigma$, 则$\eta \sim N(0,1)$. 若$\eta = -\xi$, 则$\eta \sim N(-a, \sigma^2)$.

例3 设$\xi \sim N(a, \sigma^2)$, 求$\eta = k\xi + b$的密度函数($k \neq 0$).

解: $y = kx + b$是严格单调的, 且$x = f^{-1}(y) = (y-b)/k$有连续导函数. 又ξ的密度函数为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty.$$

所以利用前面的定理可知, η的密度函数为

$$\begin{aligned} g(y) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(\frac{y-b}{k}-a)^2}{2\sigma^2}\right\} \cdot \left|\frac{1}{k}\right| \\ &= \frac{1}{\sqrt{2\pi}|k|\sigma} \exp\left\{-\frac{(y-ka-b)^2}{2k^2\sigma^2}\right\}, \quad -\infty < y < \infty. \end{aligned}$$

即, $\eta \sim N(ka+b, k^2\sigma^2)$.

例3 设$\xi \sim N(a,\sigma^2)$, 求$\eta = k\xi + b$的密度函数$(k \neq 0)$.

解: $y = kx + b$是严格单调的, 且$x = f^{-1}(y) = (y-b)/k$有连续导函数. 又ξ的密度函数为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty.$$

所以利用前面的定理可知, η的密度函数为

$$\begin{aligned} g(y) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(\frac{y-b}{k}-a)^2}{2\sigma^2}\right\} \cdot \left|\frac{1}{k}\right| \\ &= \frac{1}{\sqrt{2\pi}|k|\sigma} \exp\left\{-\frac{(y-ka-b)^2}{2k^2\sigma^2}\right\}, \quad -\infty < y < \infty. \end{aligned}$$

即, $\eta \sim N(ka+b, k^2\sigma^2)$.

注: 若$\eta = (\xi - a)/\sigma$, 则$\eta \sim N(0,1)$. 若$\eta = -\xi$, 则$\eta \sim N(-a, \sigma^2)$.

例4 $\xi \sim N(0,1)$, 求$\eta = \xi^2$的密度函数.

例4 $\xi \sim N(0,1)$, 求$\eta = \xi^2$的密度函数.

解: $y = x^2$是分段单调的. 它的反函数为:

$$x = \begin{cases} -\sqrt{y}, & x < 0, \\ \sqrt{y}, & x \geq 0. \end{cases}$$

例4 $\xi \sim N(0,1)$, 求$\eta = \xi^2$的密度函数.

解: $y = x^2$是分段单调的. 它的反函数为:

$$x = \begin{cases} -\sqrt{y}, & x < 0, \\ \sqrt{y}, & x \geq 0. \end{cases}$$

所以由前面的推论可知, 当$y > 0$时,

$$\begin{aligned} g(y) = & \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(-\sqrt{y})^2}{2} \right\} \cdot \left| \frac{-1}{2\sqrt{y}} \right| \\ & + \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(\sqrt{y})^2}{2} \right\} \cdot \left| \frac{1}{2\sqrt{y}} \right| \\ = & \frac{1}{\sqrt{2\pi y}} \exp\{-\frac{y}{2}\}. \end{aligned}$$

例4 $\xi \sim N(0,1)$, 求$\eta = \xi^2$的密度函数.

解: $y = x^2$是分段单调的. 它的反函数为:

$$x = \begin{cases} -\sqrt{y}, & x < 0, \\ \sqrt{y}, & x \geq 0. \end{cases}$$

所以由前面的推论可知, 当$y > 0$时,

$$\begin{aligned} g(y) =& \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(-\sqrt{y})^2}{2}\right\} \cdot \left|\frac{-1}{2\sqrt{y}}\right| \\ &+ \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(\sqrt{y})^2}{2}\right\} \cdot \left|\frac{1}{2\sqrt{y}}\right| \\ =& \frac{1}{\sqrt{2\pi y}} \exp\{-\frac{y}{2}\}. \end{aligned}$$

当$y \leq 0$时, $g(y) = 0$.

例4 $\xi \sim N(0,1)$, 求$\eta = \xi^2$的密度函数.

解: $y = x^2$是分段单调的. 它的反函数为:

$$x = \begin{cases} -\sqrt{y}, & x < 0, \\ \sqrt{y}, & x \geq 0. \end{cases}$$

所以由前面的推论可知, 当$y > 0$时,

$$\begin{aligned} g(y) =& \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(-\sqrt{y})^2}{2} \right\} \cdot \left| \frac{-1}{2\sqrt{y}} \right| \\ &+ \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(\sqrt{y})^2}{2} \right\} \cdot \left| \frac{1}{2\sqrt{y}} \right| \\ =& \frac{1}{\sqrt{2\pi y}} \exp\{-\frac{y}{2}\}. \end{aligned}$$

当$y \leq 0$时, $g(y) = 0$. 综上所述,

$$g(y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} \exp\{-\frac{y}{2}\}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

例5 设ξ有连续的分布函数$F(x)$, 求$\theta = F(\xi)$的分布.

例5 设ξ有连续的分布函数$F(x)$, 求$\theta = F(\xi)$的分布.

解: 因为$0 \le F(\cdot) \le 1$. 所以当$y < 0$时, $\mathsf{P}(\theta \le y) = 0$; 当$y > 1$时, $\mathsf{P}(\theta \le y) = 1$;

例5 设ξ有连续的分布函数$F(x)$，求$\theta = F(\xi)$的分布.

解：因为$0 \le F(\cdot) \le 1$. 所以当$y < 0$时，$\mathrm{P}(\theta \le y) = 0$；当$y > 1$时，$\mathrm{P}(\theta \le y) = 1$；当$0 \le y \le 1$时，我们先来定义$F(x)$的广义反函数(因为$y = F(x)$不一定严格单调递增)：

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

例5 设ξ有连续的分布函数$F(x)$, 求$\theta = F(\xi)$的分布.

解: 因为$0 \le F(\cdot) \le 1$. 所以当$y < 0$时, $\mathsf{P}(\theta \le y) = 0$; 当$y > 1$时, $\mathsf{P}(\theta \le y) = 1$; 当$0 \le y \le 1$时, 我们先来定义$F(x)$的广义反函数(因为$y = F(x)$不一定严格单调递增):

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

然后, 可知$F(F^{-1}(y)) = y$. 这是因为对任意的$\varepsilon > 0$,

$$F(F^{-1}(y) - \varepsilon) < y \le F(F^{-1}(y) + \varepsilon).$$

令$\varepsilon \searrow 0$, 由分布函数的连续性假设马上可得$F(F^{-1}(y)) = y$. 接下来我们分两种情况讨论:

(1) 若$F(x)$是严格单调递增的, 那么这里的广义反函数就是普通的反函数, 因此

$\mathsf{P}(\theta \le y) =$

例5 设ξ有连续的分布函数$F(x)$, 求$\theta = F(\xi)$的分布.

解: 因为$0 \leq F(\cdot) \leq 1$. 所以当$y < 0$时, $\mathsf{P}(\theta \leq y) = 0$; 当$y > 1$时, $\mathsf{P}(\theta \leq y) = 1$; 当$0 \leq y \leq 1$时, 我们先来定义$F(x)$的广义反函数(因为$y = F(x)$不一定严格单调递增):

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

然后, 可知$F(F^{-1}(y)) = y$. 这是因为对任意的$\varepsilon > 0$,

$$F(F^{-1}(y) - \varepsilon) < y \leq F(F^{-1}(y) + \varepsilon).$$

令$\varepsilon \searrow 0$, 由分布函数的连续性假设马上可得$F(F^{-1}(y)) = y$. 接下来我们分两种情况讨论:

(1) 若$F(x)$是严格单调递增的, 那么这里的广义反函数就是普通的反函数, 因此

$$\mathsf{P}(\theta \leq y) = \mathsf{P}(F(\xi) \leq y) = \mathsf{P}(\xi \leq F^{-1}(y)) = F(F^{-1}(y)) = y.$$

(2) 若$F(x)$不是严格单调递增的, 我们不妨假设当$F^{-1}(y) \le x \le F^{-1}(y) + a$ ($a > 0$为某常数)时都有$F(x) = y$, 且当$x > F^{-1}(y)$ $+a$时都有$F(x) > y$. 那么有

$$\begin{aligned} \mathsf{P}(\theta \le y) =& \mathsf{P}(F(\xi) \le y) \\ =& \mathsf{P}(\xi \le F^{-1}(y) + a) \\ =& F(F^{-1}(y) + a) = y. \end{aligned}$$

综上所述, $\theta = F(\xi) \sim U[0,1]$ (与ξ的具体分布无关, 只需ξ具有连续的分布函数).

(2) 若$F(x)$不是严格单调递增的, 我们不妨假设当$F^{-1}(y) \le x \le F^{-1}(y) + a$ ($a > 0$为某常数)时都有$F(x) = y$, 且当$x > F^{-1}(y) + a$时都有$F(x) > y$. 那么有

$$
\begin{aligned}
\mathsf{P}(\theta \le y) =& \mathsf{P}(F(\xi) \le y) \\
=& \mathsf{P}(\xi \le F^{-1}(y) + a) \\
=& F(F^{-1}(y) + a) = y.
\end{aligned}
$$

(2) 若$F(x)$不是严格单调递增的, 我们不妨假设当$F^{-1}(y) \le x \le F^{-1}(y) + a$ ($a > 0$为某常数)时都有$F(x) = y$, 且当$x > F^{-1}(y) + a$时都有$F(x) > y$. 那么有

$$\begin{aligned} \mathsf{P}(\theta \le y) =& \mathsf{P}(F(\xi) \le y) \\ =& \mathsf{P}(\xi \le F^{-1}(y) + a) \\ =& F(F^{-1}(y) + a) = y. \end{aligned}$$

综上所述, $\theta = F(\xi) \sim U[0,1]$ (与ξ的具体分布无关, 只需ξ具有连续的分布函数).

例6 若$\theta \sim U[0,1]$, $F(x)$满足分布函数的三个性质. 求$\xi = F^{-1}(\theta)$的分布, 其中$F^{-1}(y) = \sup\{x : F(x) < y\}$.

例6 若$\theta \sim U[0,1]$, $F(x)$满足分布函数的三个性质. 求$\xi = F^{-1}(\theta)$的分布, 其中$F^{-1}(y) = \sup\{x : F(x) < y\}$.

解: 首先我们来证明

$$\{F(x) \geq y\} = \{x \geq F^{-1}(y)\}. \tag{2.6.1}$$

例6 若$\theta \sim U[0,1]$, $F(x)$满足分布函数的三个性质. 求$\xi = F^{-1}(\theta)$的分布, 其中$F^{-1}(y) = \sup\{x : F(x) < y\}$.

解: 首先我们来证明

$$\{F(x) \geq y\} = \{x \geq F^{-1}(y)\}. \tag{2.6.1}$$

因为由$F^{-1}(\cdot)$的定义,

$$\begin{aligned}
x > F^{-1}(y) &\Longrightarrow F(x) \geq y; \\
x = F^{-1}(y) &\Longrightarrow x + \varepsilon > F^{-1}(y), \forall \varepsilon > 0 \\
&\Longrightarrow F(x+\varepsilon) \geq y, \forall \varepsilon > 0 \\
&\Longrightarrow F(x) = \lim_{\varepsilon \searrow 0} F(x+\varepsilon) \geq y.
\end{aligned}$$

所以

$$\{x \geq F^{-1}(y)\} \subset \{F(x) \geq y\}. \tag{2.6.2}$$

另一方面,

$$x < F^{-1}(y) \Longrightarrow F(x) < y.$$

所以,

$$\{F(x) \geq y\} \subset \{x \geq F^{-1}(y)\}. \tag{2.6.3}$$

另一方面,

$$x < F^{-1}(y) \Longrightarrow F(x) < y.$$

所以,

$$\{F(x) \geq y\} \subset \{x \geq F^{-1}(y)\}. \tag{2.6.3}$$

(2.6.2)+(2.6.3)$\Longrightarrow$(2.6.1).

另一方面,

$$x < F^{-1}(y) \Longrightarrow F(x) < y.$$

所以,

$$\{F(x) \geq y\} \subset \{x \geq F^{-1}(y)\}. \tag{2.6.3}$$

(2.6.2)+(2.6.3)⟹(2.6.1). 所以ξ的分布函数为

$$\mathsf{P}(\xi \leq x) = \mathsf{P}(F^{-1}(\theta) \leq x) = \mathsf{P}(\theta \leq F(x)) = F(x), \quad x \in \mathbb{R}.$$

另一方面,

$$x < F^{-1}(y) \Longrightarrow F(x) < y.$$

所以,

$$\{F(x) \geq y\} \subset \{x \geq F^{-1}(y)\}. \tag{2.6.3}$$

(2.6.2)+(2.6.3)$\Longrightarrow$(2.6.1). 所以ξ的分布函数为

$$\mathsf{P}(\xi \leq x) = \mathsf{P}(F^{-1}(\theta) \leq x) = \mathsf{P}(\theta \leq F(x)) = F(x), \quad x \in \mathbb{R}.$$

注: (1)这个例子提供了随机变量的存在性证明. 给定一个满足分布函数的三个性质的函数$F(x)$, 则可构造一个随机变量ξ, 使得ξ的分布函数为$F(x)$. (2)这个例子也提供了一种产生来自$F(x)$分布的随机数的方法.

2.6.3 (连续型)随机向量的函数

2.6.3 (连续型)随机向量的函数

设$(\xi_1, \cdots, \xi_n)$为连续型随机向量, 其密度函数为$p(x_1, \cdots, x_n)$.
又设$\eta = f(\xi_1, \cdots, \xi_n)$. 则$\eta$的分布函数为

$$\begin{aligned} F_\eta(y) =& \mathsf{P}(f(\xi_1, \cdots, \xi_n) \le y) \\ =& \int \cdots \int_{f(x_1, \cdots, x_n) \le y} p(x_1, \cdots, x_n) \mathrm{d}x_1 \cdots \mathrm{d}x_n. \end{aligned}$$

一些特殊的f函数:

一些特殊的f函数:

1. $\eta = \xi_1 + \xi_2$, (ξ_1, ξ_2)为连续型随机向量.

一些特殊的f函数:

1. $\eta=\xi_1+\xi_2$, (ξ_1,ξ_2)为连续型随机向量.

η的分布函数为

$$
\begin{aligned}
F_\eta(y) &= \iint_{x_1+x_2\le y} p(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2 \\
&=
\end{aligned}
$$

一些特殊的f函数:

1. $\eta=\xi_1+\xi_2$, (ξ_1,ξ_2)为连续型随机向量.

η的分布函数为

$$
\begin{aligned}
F_\eta(y) &= \iint_{x_1+x_2\le y} p(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2 \\
&= \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{y-x_1} p(x_1,x_2)\mathrm{d}x_2\Big)\mathrm{d}x_1 \\
&\overset{x_2=z-x_1}{=} \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{y} p(x_1,z-x_1)\mathrm{d}z\Big)\mathrm{d}x_1 \\
&= \int_{-\infty}^{y}\Big(\int_{-\infty}^{\infty} p(x_1,z-x_1)\mathrm{d}x_1\Big)\mathrm{d}z.
\end{aligned}
$$

因此η是连续型随机变量, 且其密度函数为

$$p_\eta(z) = \int_{-\infty}^{\infty} p(x, z - x)\mathrm{d}x.$$

因此η是连续型随机变量, 且其密度函数为

$$p_\eta(z) = \int_{-\infty}^{\infty} p(x, z-x)\mathrm{d}x.$$

同理可得

$$p_\eta(z) = \int_{-\infty}^{\infty} p(z-y, y)\mathrm{d}y.$$

因此η是连续型随机变量, 且其密度函数为

$$p_\eta(z) = \int_{-\infty}^{\infty} p(x, z-x)\mathrm{d}x.$$

同理可得

$$p_\eta(z) = \int_{-\infty}^{\infty} p(z-y, y)\mathrm{d}y.$$

若ξ_1与ξ_2相互独立, 则有

$$\begin{aligned} p_\eta(z) &= \int_{-\infty}^{\infty} p_{\xi_1}(x)p_{\xi_2}(z-x)\mathrm{d}x \\ &= \int_{-\infty}^{\infty} p_{\xi_1}(z-y)p_{\xi_2}(y)\mathrm{d}y. \end{aligned} \tag{2.6.4}$$

称(2.6.4)为(连续)卷积公式(convolution formula).

例7 设ξ与η相互独立, 都服从$N(0,1)$. 求$\zeta=\xi+\eta$的密度函数.

例7 设ξ与η相互独立, 都服从$N(0,1)$. 求$\zeta=\xi+\eta$的密度函数.

解: 利用卷积公式(2.6.4), 可知对任意的$z\in\mathbb{R}$,

$$\begin{aligned}
p_\zeta(z) &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{(z-x)^2}{2}}\mathrm{d}x\\
&= \int_{-\infty}^{\infty}\frac{1}{2\pi}e^{-\frac{z^2}{4}}e^{-(x-\frac{z}{2})^2}\mathrm{d}x\\
&= \frac{1}{\sqrt{2\pi}\sqrt{2}}e^{-\frac{z^2}{4}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\frac{1}{\sqrt{2}}}e^{-\frac{(x-\frac{z}{2})^2}{2(1/\sqrt{2})^2}}\mathrm{d}x\\
&= \frac{1}{\sqrt{2\pi}\sqrt{2}}e^{-\frac{z^2}{4}}.
\end{aligned}$$

所以, $\zeta=\xi+\eta\sim N(0,2)$.

事实上, 利用卷积公式, 我们可以得到一个更一般的结论(由于计算过程较繁琐, 这里我们不给出计算过程, 在下一章中我们将采用更简单的方法来证明):

设ξ与η相互独立, $\xi \sim N(a_1, \sigma_1^2)$, $\eta \sim N(a_2, \sigma_2^2)$. 则

$$\zeta = \xi + \eta \sim N(a_1 + a_2, \sigma_1^2 + \sigma_2^2).$$

事实上, 利用卷积公式, 我们可以得到一个更一般的结论(由于计算过程较繁琐, 这里我们不给出计算过程, 在下一章中我们将采用更简单的方法来证明):

设ξ与η相互独立, $\xi \sim N(a_1, \sigma_1^2)$, $\eta \sim N(a_2, \sigma_2^2)$. 则

$$\zeta = \xi + \eta \sim N(a_1 + a_2, \sigma_1^2 + \sigma_2^2).$$

更一般地, 若$\xi_i \sim N(a_i, \sigma_i^2), i = 1, \cdots, n$, 且$\xi_1, \cdots, \xi_n$相互独立.

则

$$\sum_{i=1}^{n} \xi_i \sim N(\sum_{i=1}^{n} a_i, \sum_{i=1}^{n} \sigma_i^2).$$

例8 设ξ与η相互独立, 密度函数分别为

$$p_\xi(x)=\begin{cases} ae^{-ax}, & x>0, \\ 0, & x\le 0, \end{cases} \quad (a>0);$$

$$p_\eta(x)=\begin{cases} be^{-bx}, & x>0, \\ 0, & x\le 0, \end{cases} \quad (b>0).$$

求$\zeta=\xi+\eta$的密度函数.

例8 设ξ与η相互独立, 密度函数分别为

$$p_\xi(x) = \begin{cases} ae^{-ax}, & x > 0, \\ 0, & x \le 0, \end{cases} \quad (a > 0);$$

$$p_\eta(x) = \begin{cases} be^{-bx}, & x > 0, \\ 0, & x \le 0, \end{cases} \quad (b > 0).$$

求$\zeta = \xi + \eta$的密度函数.

解: 当且仅当$x > 0$且$z - x > 0$时, 即$z > x > 0$时,

$$p_\xi(x) p_\eta(z - x) \neq 0.$$

例8 设ξ与η相互独立, 密度函数分别为

$$p_\xi(x)=\begin{cases} ae^{-ax}, & x>0, \\ 0, & x\le 0, \end{cases} \quad (a>0);$$

$$p_\eta(x)=\begin{cases} be^{-bx}, & x>0, \\ 0, & x\le 0, \end{cases} \quad (b>0).$$

求$\zeta=\xi+\eta$的密度函数.

解: 当且仅当$x>0$且$z-x>0$时, 即$z>x>0$时,

$$p_\xi(x)p_\eta(z-x)\neq 0.$$

因此, 由卷积公式(2.6.4)可知:
当$z\le 0$时, $p_\zeta(z)=0$;

当$z>0$时,

$$
\begin{aligned}
p_\zeta(z) &= \int_0^z ae^{-ax}be^{-b(z-x)}\mathrm{d}x \\
&= abe^{-bz}\int_0^z e^{-(a-b)x}\mathrm{d}x \\
&= \begin{cases} abze^{-bz}, & a=b, \\ \frac{ab}{a-b}(e^{-bz}-e^{-az}), & a\neq b. \end{cases}
\end{aligned}
$$

当$z > 0$时,

$$
\begin{aligned}
p_\zeta(z) &= \int_0^z ae^{-ax}be^{-b(z-x)}\mathrm{d}x \\
&= abe^{-bz}\int_0^z e^{-(a-b)x}\mathrm{d}x \\
&= \begin{cases} abze^{-bz}, & a = b, \\ \frac{ab}{a-b}(e^{-bz} - e^{-az}), & a \neq b. \end{cases}
\end{aligned}
$$

综上所述,

$$
p_\zeta(z) = \begin{cases} 0, & z \leq 0, \\ abze^{-bz}, & z > 0, \end{cases} \quad (\text{当}a = b).
$$

当$z>0$时,

$$
\begin{aligned}
p_\zeta(z) &= \int_0^z ae^{-ax}be^{-b(z-x)}\mathrm{d}x \\
&= abe^{-bz}\int_0^z e^{-(a-b)x}\mathrm{d}x \\
&= \begin{cases} abze^{-bz}, & a=b, \\ \frac{ab}{a-b}(e^{-bz}-e^{-az}), & a\neq b. \end{cases}
\end{aligned}
$$

综上所述,

$$
p_\zeta(z) = \begin{cases} 0, & z\leq 0, \\ abze^{-bz}, & z>0, \end{cases} \quad (\text{当}a=b).
$$

$$
p_\zeta(z) = \begin{cases} 0, & z\leq 0, \\ \frac{ab}{a-b}(e^{-bz}-e^{-az}), & z>0, \end{cases} \quad (\text{当}a\neq b).
$$

2. $\zeta = \xi/\eta$, (ξ, η)为连续型随机向量.

2. $\zeta=\xi/\eta$, (ξ,η)为连续型随机向量.

ζ的分布函数为

$$
\begin{aligned}
& F_\zeta(z) \\
= \quad & \mathrm{P}(\xi/\eta \le z) = \int\int_{x/y\le z} p(x,y)\mathrm{d}x\mathrm{d}y \\
= \quad &
\end{aligned}
$$

2. $\zeta=\xi/\eta$, (ξ,η)为连续型随机向量.

ζ的分布函数为

$$
\begin{aligned}
& F_\zeta(z) \\
= \; & \mathsf{P}(\xi/\eta \le z) = \int\int_{x/y\le z} p(x,y)\mathrm{d}x\mathrm{d}y \\
= \; & \int_0^\infty \Big(\int_{-\infty}^{yz} p(x,y)\mathrm{d}x\Big)\mathrm{d}y + \int_{-\infty}^0 \Big(\int_{yz}^\infty p(x,y)\mathrm{d}x\Big)\mathrm{d}y \\
\overset{x=uy}{=} \; & \int_{-\infty}^{z} \Big(\int_0^\infty p(uy,y)y\mathrm{d}y - \int_{-\infty}^0 p(uy,y)y\mathrm{d}y\Big)\mathrm{d}u \\
= \; & \int_{-\infty}^{z} \left(\int_{-\infty}^\infty p(uy,y)|y|\mathrm{d}y\right)\mathrm{d}u.
\end{aligned}
$$

这说明: 若(ξ,η)是连续型随机向量, 则$\zeta=\xi/\eta$是连续型随机变量, 其密度函数为

$$p_\zeta(z)=\int_{-\infty}^{\infty}p(zy,y)|y|\mathrm{d}y.$$

例9 设ξ与η相互独立, 都服从$U[0,a]$, 求ξ/η的密度函数.

例9 设ξ与η相互独立, 都服从$U[0,a]$, 求ξ/η的密度函数.

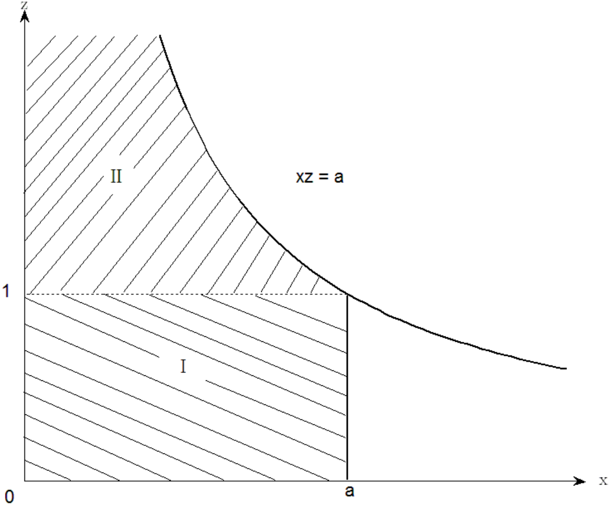
解: 易知

$$p_\xi(x) = p_\eta(x) = \begin{cases} 1/a, & 0 \le x \le a, \\ 0, & \text{其他}. \end{cases}$$

又ξ与η相互独立, 因此只有当

$$0 \le xz \le a \ \text{ 且 } \ 0 \le x \le a\text{时},$$

$$p(zx, x) = p_\xi(xz)p_\eta(x) = \frac{1}{a^2} \ne 0.$$


z
II
xz = a
1
I
0
a
x

当$z < 0$时, $p(zx, x) = 0$, 所以$p_{\xi/\eta}(z) = 0$;

当$z < 0$时, $p(zx, x) = 0$, 所以$p_{\xi/\eta}(z) = 0$;

当$0 \leq z < 1$时,

$$p_{\xi/\eta}(z) = \int_0^a \frac{1}{a^2} x \mathrm{d}x = \frac{1}{2};$$

当$z<0$时, $p(zx,x)=0$, 所以$p_{\xi/\eta}(z)=0$;

当$0\le z<1$时,

$$p_{\xi/\eta}(z)=\int_0^a \frac{1}{a^2}x\mathsf{d}x=\frac{1}{2};$$

当$1\le z<\infty$时,

$$p_{\xi/\eta}(z)=\int_0^{a/z} \frac{1}{a^2}x\mathsf{d}x=\frac{1}{2z^2}.$$

当$z<0$时，$p(zx,x)=0$，所以$p_{\xi/\eta}(z)=0$;

当$0\le z<1$时，

$$p_{\xi/\eta}(z)=\int_0^a \frac{1}{a^2}x\mathrm{d}x=\frac{1}{2};$$

当$1\le z<\infty$时，

$$p_{\xi/\eta}(z)=\int_0^{a/z}\frac{1}{a^2}x\mathrm{d}x=\frac{1}{2z^2}.$$

综上所述，

$$p_{\xi/\eta}(z)=\begin{cases}\frac{1}{2}, & 0\le z<1,\\ \frac{1}{2z^2}, & 1\le z<\infty,\\ 0, & \text{其他}.\end{cases}$$

3. 次序统计量(order statistic)的分布.

3. 次序统计量(order statistic)的分布.

设$\xi_1, \xi_2, \cdots, \xi_n$独立同分布, 分布函数为$F(x)$. 把$\xi_1, \xi_2, \cdots, \xi_n$的每一组取值$\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega)(\omega \in \Omega)$都按大小次序排列, 得到随机变量$\xi_1^*, \xi_2^*, \cdots, \xi_n^*$, 称为次序统计量.

3. 次序统计量(order statistic)的分布.

设$\xi_1, \xi_2, \cdots, \xi_n$独立同分布, 分布函数为$F(x)$. 把$\xi_1, \xi_2, \cdots, \xi_n$的每一组取值$\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega)(\omega \in \Omega)$都按大小次序排列, 得到随机变量$\xi_1^*, \xi_2^*, \cdots, \xi_n^*$, 称为次序统计量. 次序统计量满足

$$\xi_1^* \le \xi_2^* \le \cdots \le \xi_n^*.$$

3. 次序统计量(order statistic)的分布.

设$\xi_1, \xi_2, \cdots, \xi_n$独立同分布, 分布函数为$F(x)$. 把$\xi_1, \xi_2, \cdots, \xi_n$的每一组取值$\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega)(\omega \in \Omega)$都按大小次序排列, 得到随机变量$\xi_1^*, \xi_2^*, \cdots, \xi_n^*$, 称为次序统计量. 次序统计量满足

$$\xi_1^* \le \xi_2^* \le \cdots \le \xi_n^*.$$

且由定义,

$$\xi_1^* = \min\{\xi_1, \xi_2, \cdots, \xi_n\}, \quad \xi_n^* = \max\{\xi_1, \xi_2, \cdots, \xi_n\}.$$

我们来求ξ_1^*, ξ_n^*以及(ξ_1^*, ξ_n^*)的分布, 这在数理统计中是有用的.

我们来求ξ_1^*, ξ_n^*以及(ξ_1^*, ξ_n^*)的分布，这在数理统计中是有用的.

(1) ξ_n^*的分布函数

我们来求ξ_1^*, ξ_n^*以及(ξ_1^*, ξ_n^*)的分布, 这在数理统计中是有用的.

(1) ξ_n^*的分布函数

$$
\begin{aligned}
\mathrm{P}(\xi_n^* \le x) =& \mathrm{P}(\xi_1 \le x, \xi_2 \le x, \cdots, \xi_n \le x) \\
=& \mathrm{P}(\xi_1 \le x)\mathrm{P}(\xi_2 \le x)\cdots\mathrm{P}(\xi_n \le x) \\
=& [F(x)]^n.
\end{aligned}
$$

我们来求ξ_1^*, ξ_n^*以及(ξ_1^*, ξ_n^*)的分布，这在数理统计中是有用的.

(1) ξ_n^*的分布函数

$$
\begin{aligned}
\mathsf{P}(\xi_n^* \leq x) =& \mathsf{P}(\xi_1 \leq x, \xi_2 \leq x, \cdots, \xi_n \leq x) \\
=& \mathsf{P}(\xi_1 \leq x)\mathsf{P}(\xi_2 \leq x) \cdots \mathsf{P}(\xi_n \leq x) \\
=& [F(x)]^n.
\end{aligned}
$$

(2) ξ_1^*的分布函数

我们来求ξ_1^*, ξ_n^*以及(ξ_1^*, ξ_n^*)的分布，这在数理统计中是有用的.

(1) ξ_n^*的分布函数

$$\begin{aligned}
\mathsf{P}(\xi_n^* \le x) =& \mathsf{P}(\xi_1 \le x, \xi_2 \le x, \cdots, \xi_n \le x) \\
=& \mathsf{P}(\xi_1 \le x)\mathsf{P}(\xi_2 \le x)\cdots\mathsf{P}(\xi_n \le x) \\
=& [F(x)]^n.
\end{aligned}$$

(2) ξ_1^*的分布函数

$$\begin{aligned}
\mathsf{P}(\xi_1^* > x) =& \mathsf{P}(\xi_1 > x, \xi_2 > x, \cdots, \xi_n > x) \\
=& \mathsf{P}(\xi_1 > x)\mathsf{P}(\xi_2 > x)\cdots\mathsf{P}(\xi_n > x) \\
=& [1 - F(x)]^n,
\end{aligned}$$

我们来求ξ_1^*, ξ_n^*以及(ξ_1^*, ξ_n^*)的分布，这在数理统计中是有用的.

(1) ξ_n^*的分布函数

$$\begin{aligned}\mathsf{P}(\xi_n^* \le x) =& \mathsf{P}(\xi_1 \le x, \xi_2 \le x, \cdots, \xi_n \le x) \\ =& \mathsf{P}(\xi_1 \le x)\mathsf{P}(\xi_2 \le x)\cdots\mathsf{P}(\xi_n \le x) \\ =& [F(x)]^n.\end{aligned}$$

(2) ξ_1^*的分布函数

$$\begin{aligned}\mathsf{P}(\xi_1^* > x) =& \mathsf{P}(\xi_1 > x, \xi_2 > x, \cdots, \xi_n > x) \\ =& \mathsf{P}(\xi_1 > x)\mathsf{P}(\xi_2 > x)\cdots\mathsf{P}(\xi_n > x) \\ =& [1 - F(x)]^n,\end{aligned}$$

所以

$$\mathsf{P}(\xi_1^* \le x) = 1 - [1 - F(x)]^n.$$

(3) (ξ_1^*, ξ_n^*)的分布函数

(3) (ξ_1^*, ξ_n^*)的分布函数

$$
\begin{aligned}
F(x,y) =& \mathsf{P}(\xi_1^* \leq x, \xi_n^* \leq y) \\
=& \mathsf{P}(\xi_n^* \leq y) - \mathsf{P}(\xi_1^* > x, \xi_n^* \leq y) \\
=& [F(y)]^n - \mathsf{P}(\bigcap_{i=1}^{n} \{x < \xi_i \leq y\}) \\
=& \begin{cases} [F(y)]^n - [F(y) - F(x)]^n, & x < y, \\ [F(y)]^n, & x \geq y. \end{cases}
\end{aligned}
$$

2.6.4 随机向量的变换

2.6.4 随机向量的变换

设$(\xi_1,\cdots,\xi_n)$的密度函数为$p(x_1,\cdots,x_n)$. 现有m个Borel函数:

$$\eta_1=f_1(\xi_1,\cdots,\xi_n),\cdots,\eta_m=f_m(\xi_1,\cdots,\xi_n),$$

则$(\eta_1,\cdots,\eta_m)$是随机向量, 其联合分布函数为

$$\begin{aligned} & G(y_1,\cdots,y_m) \\ =& \mathrm{P}(\eta_1\le y_1,\cdots,\eta_m\le y_m) \\ =& \int\cdots\int_D p(x_1,\cdots,x_n)dx_1\cdots dx_n, \end{aligned} \tag{2.6.5}$$

这里, D是n维区域(n维Borel集):

$$\{(x_1,\cdots,x_n): f_1(x_1,\cdots,x_n)\le y_1,\cdots,f_m(x_1,\cdots,x_n)\le y_m\}.$$

关于$(\eta_1, \cdots, \eta_m)$的联合密度函数, 我们有下列的结论:

关于$(\eta_1, \cdots, \eta_m)$的联合密度函数, 我们有下列的结论:

定理

如果$m = n$, $\{u_j = f_j(x_1, \cdots, x_n), j = 1, \cdots, n\}$有唯一的反函数组:

$$x_i = x_i(u_1, \cdots, u_n), i = 1, \cdots, n,$$

且

$$J = \frac{\partial(x_1, \cdots, x_n)}{\partial(u_1, \cdots, u_n)} \neq 0,$$

则$(\eta_1, \cdots, \eta_n)$是连续型随机向量. 当$(y_1, \cdots, y_n) \in (f_1, \cdots, f_n)$的值域时, 其密度函数为

$$q(u_1, \cdots, u_n) = p(x_1(u_1, \cdots, u_n), \cdots, x_n(u_1, \cdots, u_n)) \cdot |J| \quad (2.6.6)$$

其它情况$q(u_1, \cdots, u_n) = 0$.

证明: 在(2.6.5)中利用重积分的变量替换:

$$u_1 = f_1(x_1, \cdots, x_n), \cdots, u_n = f_n(x_1, \cdots, x_n).$$

当$(y_1, \cdots, y_n) \notin (f_1, \cdots, f_n)$的值域时, $G(y_1, \cdots, y_n)$要么为0要么为1. 当$(y_1, \cdots, y_n) \in (f_1, \cdots, f_n)$ 的值域时, 由(2.6.5)得$(\eta_1, \cdots, \eta_n)$的联合分布函数为

$$\begin{aligned} & G(y_1, \cdots, y_n) \\ = & \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_n} p(x_1(u_1, \cdots, u_n), \cdots, x_n(u_1, \cdots, u_n)) \cdot |J| \mathrm{d}u_1 \cdots \mathrm{d}u_n, \end{aligned}$$

所以$(\eta_1, \cdots, \eta_n)$是连续型随机向量, 且$q(u_1, \cdots, u_n)$为$(\eta_1, \cdots, \eta_n)$的联合密度函数.

例10 设ξ,η相互独立, 都服从参数为1的指数分布. 求:
(1) $\alpha=\xi+\eta$与$\beta=\xi/\eta$的联合密度函数;
(2) 分别求$\xi+\eta$与ξ/η的边际密度函数.

例10 设ξ, η相互独立, 都服从参数为1的指数分布. 求:
(1) $\alpha = \xi + \eta$与$\beta = \xi / \eta$的联合密度函数;
(2) 分别求$\xi + \eta$与ξ / η的边际密度函数.

解: (1) (ξ, η)的联合密度函数为

$$p(x, y) = \begin{cases} e^{-(x+y)}, & x > 0, y > 0, \\ 0, & \text{其他}. \end{cases}$$

例10 设ξ,η相互独立，都服从参数为1的指数分布．求:
(1) $\alpha=\xi+\eta$与$\beta=\xi/\eta$的联合密度函数;
(2) 分别求$\xi+\eta$与ξ/η的边际密度函数.

解: (1) (ξ,η)的联合密度函数为

$$p(x,y)=\begin{cases} e^{-(x+y)}, & x>0, y>0, \\ 0, & \text{其他}. \end{cases}$$

函数组

$$\begin{cases} u=x+y \\ v=x/y \end{cases}$$

有唯一的反函数组:

$$\begin{cases} x=uv/(1+v) \\ y=u/(1+v) \end{cases}.$$

(u,v)的值域为$\{u>0, v>0\}$.

$$J=\frac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix} \frac{v}{1+v}, & \frac{u}{(1+v)^2} \\ \frac{1}{1+v}, & -\frac{u}{(1+v)^2} \end{vmatrix}=-\frac{u}{(1+v)^2},$$

所以

$$|J|=\frac{u}{(1+v)^2}.$$

根据(2.6.6), 容易求得(α,β)的联合密度函数为

$$q(u,v)=\begin{cases} \frac{ue^{-u}}{(1+v)^2}, & u>0, v>0, \\ 0, & \text{其他}. \end{cases}$$

(2) 从联合密度函数可以求得两个边际密度函数分别为

$$p_\alpha(u) = \begin{cases} ue^{-u}, & u > 0, \\ 0, & u \leq 0, \end{cases}$$

$$p_\beta(v) = \begin{cases} \frac{1}{(1+v)^2}, & v > 0, \\ 0, & v \leq 0, \end{cases}$$

可以看出, α与β相互独立.

这个例子告诉我们:

- 要判断随机向量的几个函数$(\eta_1, \cdots, \eta_n)$是否独立, 可用随机向量变换公式求得它们的联合分布, 以及相应的边际分布, 然后利用独立性的各种充要条件来判断.
- 为求随机向量的一个函数的分布, 有时可以适当补充几个函数, 先求它们的联合分布, 而原来要求的随机向量的函数的分布可以作为其边际分布得到.

例11 ξ,η相互独立, 并且$\xi\sim N(0,\sigma^2)$, $\eta\sim U(0,\pi)$. 求$\alpha=\xi+a\cos\eta$的密度函数, 其中a为常数.

例11 ξ,η相互独立，并且$\xi\sim N(0,\sigma^2)$，$\eta\sim U(0,\pi)$．求$\alpha=\xi+a\cos\eta$的密度函数，其中a为常数.

解：(ξ,η)的联合密度函数为

$$p(x,y)=\begin{cases}\frac{1}{\sqrt{2\pi}\sigma\pi}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}, & -\infty<x<\infty, 0<y<\pi,\\ 0, & \text{其他}.\end{cases}$$

例11 ξ, η相互独立, 并且$\xi \sim N(0, \sigma^2)$, $\eta \sim U(0, \pi)$. 求$\alpha = \xi + a\cos\eta$的密度函数, 其中a为常数.

解: (ξ, η)的联合密度函数为

$$p(x,y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma\pi} \exp\left\{ -\frac{x^2}{2\sigma^2} \right\}, & -\infty < x < \infty, 0 < y < \pi, \\ 0, & \text{其他}. \end{cases}$$

令$\beta = \eta$, 那么与变换

$$\begin{cases} \alpha = \xi + a\cos\eta \\ \beta = \eta \end{cases}$$

相对应的函数组是

$$\begin{cases} u = x + a\cos y \\ v = y \end{cases}.$$

它有唯一解:

$$\begin{cases} x = u - a\cos v \\ y = v \end{cases}.$$

它有唯一解:

$$\begin{cases} x = u - a\cos v \\ y = v \end{cases}.$$

(u,v)的值域是$\{u \in \mathbb{R}, 0 < v < \pi\}$. 所以可得$|J| = 1$, (u,v)属于它的值域时

$$q(u,v) = p(u - a\cos v, v),$$

(u,v)不属于它的值域时, $q(u,v) = 0$.

它有唯一解:

$$\begin{cases} x = u - a\cos v \\ y = v \end{cases}.$$

(u,v)的值域是$\{u\in\mathbb{R}, 0<v<\pi\}$. 所以可得$|J|=1$, (u,v)属于它的值域时

$$q(u,v) = p(u - a\cos v, v),$$

(u,v)不属于它的值域时, $q(u,v)=0$. 因此, α的密度函数为

$$\begin{aligned} p_\alpha(u) &= \int_{-\infty}^{\infty} q(u,v)\mathrm{d}v \\ &= \int_0^{\pi} p(u - a\cos v, v)\mathrm{d}v \\ &= \frac{1}{\sqrt{2\pi}\sigma\pi}\int_0^{\pi} \exp\Big\{-\frac{(u-a\cos v)^2}{2\sigma^2}\Big\}\mathrm{d}v, \quad u\in\mathbb{R}. \end{aligned}$$

例12 ξ与η独立同分布，都服从$N(0,1)$分布，$\xi=\rho\cos\phi$, $\eta=\rho\sin\phi$.
求证：$\rho=\rho(\xi,\eta)$与$\phi=\phi(\xi,\eta)$相互独立.

例12 ξ与η独立同分布, 都服从$N(0,1)$分布, $\xi=\rho\cos\phi$, $\eta=\rho\sin\phi$. 求证: $\rho=\rho(\xi,\eta)$与$\phi=\phi(\xi,\eta)$相互独立.

解: 先求(ρ,ϕ)的联合密度函数. 变换的函数组与反函数组分别为

$$\left\{\begin{array}{l} r=\rho(x,y) \\ \theta=\phi(x,y) \end{array}\right., \quad \left\{\begin{array}{l} x=r\cos\theta \\ y=r\sin\theta \end{array}\right..$$

例12 ξ与η独立同分布，都服从$N(0,1)$分布，$\xi=\rho\cos\phi$，$\eta=\rho\sin\phi$。求证：$\rho=\rho(\xi,\eta)$与$\phi=\phi(\xi,\eta)$相互独立.

解：先求(ρ,ϕ)的联合密度函数. 变换的函数组与反函数组分别为

$$\begin{cases} r=\rho(x,y) \\ \theta=\phi(x,y) \end{cases}, \qquad \begin{cases} x=r\cos\theta \\ y=r\sin\theta \end{cases}.$$

这个变换在$\{-\infty<x<\infty,-\infty<y<\infty,(x,y)\neq(0,0)\}$与$\{r>0,0\leq\theta<2\pi\}$间是一一对应的.

$$J=\begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}=r.$$

而(ξ,η)的联合密度函数为

$$p(x,y)=\frac{1}{2\pi}\exp\Big\{-\frac{x^2+y^2}{2}\Big\},\quad x,y\in\mathbb{R}.$$

所以, (ρ,ϕ)的联合密度函数为

$$q(r,\theta)=\begin{cases}\frac{r}{2\pi}e^{-\frac{r^2}{2}}, & r>0, 0\le\theta<2\pi,\\ 0, & \text{其他}.\end{cases}$$

所以, (ρ,ϕ)的联合密度函数为

$$q(r,\theta)=\begin{cases}\frac{r}{2\pi}e^{-\frac{r^2}{2}}, & r>0, 0\le\theta<2\pi,\\ 0, & \text{其他}.\end{cases}$$

容易求得ρ和ϕ的边际密度函数分别为

$$\begin{cases}re^{-\frac{r^2}{2}}, & r>0,\\ 0, & \text{其他},\end{cases}\qquad \begin{cases}\frac{1}{2\pi}, & 0\le\theta<2\pi,\\ 0, & \text{其他}.\end{cases}$$

所以, (ρ, ϕ)的联合密度函数为

$$q(r,\theta) = \begin{cases} \frac{r}{2\pi} e^{-\frac{r^2}{2}}, & r > 0, 0 \le \theta < 2\pi, \\ 0, & \text{其他}. \end{cases}$$

容易求得ρ和ϕ的边际密度函数分别为

$$\begin{cases} re^{-\frac{r^2}{2}}, & r > 0, \\ 0, & \text{其他}, \end{cases} \qquad \begin{cases} \frac{1}{2\pi}, & 0 \le \theta < 2\pi, \\ 0, & \text{其他}. \end{cases}$$

所以ρ与ϕ相互独立. 称ρ的分布为瑞利(Rayleigh)分布.

例13 设ξ_1,ξ_2相互独立, 都服从$U(0,1)$分布.

$$\eta_1=(-2\ln\xi_1)^{1/2}\cos(2\pi\xi_2),\quad \eta_2=(-2\ln\xi_1)^{1/2}\sin(2\pi\xi_2).$$

证明η_1,η_2相互独立, 都服从$N(0,1)$分布.

例13 设ξ_1,ξ_2相互独立, 都服从$U(0,1)$分布.

$$\eta_1=(-2\ln\xi_1)^{1/2}\cos(2\pi\xi_2),\quad \eta_2=(-2\ln\xi_1)^{1/2}\sin(2\pi\xi_2).$$

证明η_1,η_2相互独立, 都服从$N(0,1)$分布.

解: (ξ_1,ξ_2)的联合密度函数为

$$p(x,y)=\begin{cases}1, & 0<x<1, 0<y<1,\\ 0, & 其他.\end{cases}$$

记

$$\begin{cases}u=(-2\ln x)^{1/2}\cos(2\pi y),\\ v=(-2\ln x)^{1/2}\sin(2\pi y).\end{cases}$$

例13 设ξ_1,ξ_2相互独立, 都服从$U(0,1)$分布.

$$\eta_1=(-2\ln\xi_1)^{1/2}\cos(2\pi\xi_2),\quad \eta_2=(-2\ln\xi_1)^{1/2}\sin(2\pi\xi_2).$$

证明η_1,η_2相互独立, 都服从$N(0,1)$分布.

解: (ξ_1,ξ_2)的联合密度函数为

$$p(x,y)=\begin{cases}1, & 0<x<1,0<y<1,\\ 0, & \text{其他}.\end{cases}$$

记

$$\begin{cases}u=(-2\ln x)^{1/2}\cos(2\pi y),\\ v=(-2\ln x)^{1/2}\sin(2\pi y).\end{cases}$$

容易看出u,v的值域均为$\mathbb{R}$, 且上述方程组的反函数组为

$$\begin{cases}x=\exp\left\{-\frac{u^2+v^2}{2}\right\},\\ y=\frac{1}{2\pi}\arctan(v/u).\end{cases}$$

因此

$$J = \begin{vmatrix} -u\exp\left\{-\frac{u^2+v^2}{2}\right\} & -v\exp\left\{-\frac{u^2+v^2}{2}\right\} \\ -\frac{1}{2\pi}\frac{v}{u^2+v^2} & \frac{1}{2\pi}\frac{u}{u^2+v^2} \end{vmatrix}$$
$$= -\frac{1}{2\pi}\exp\left\{-\frac{u^2+v^2}{2}\right\}.$$

因此

$$
\begin{aligned}
J =& \begin{vmatrix} -u\exp\left\{-\frac{u^2+v^2}{2}\right\} & -v\exp\left\{-\frac{u^2+v^2}{2}\right\} \\ -\frac{1}{2\pi}\frac{v}{u^2+v^2} & \frac{1}{2\pi}\frac{u}{u^2+v^2} \end{vmatrix} \\
=& -\frac{1}{2\pi}\exp\left\{-\frac{u^2+v^2}{2}\right\}.
\end{aligned}
$$

所以，(η_1, η_2)的联合密度函数为

$$
\begin{aligned}
q(u,v) =& p(x(u,v), y(u,v))|J| \\
=& \frac{1}{2\pi}\exp\left\{-\frac{u^2+v^2}{2}\right\}, \quad u, v \in \mathbb{R}.
\end{aligned}
$$

因此

$$J = \begin{vmatrix} -u\exp\{-\frac{u^2+v^2}{2}\} & -v\exp\{-\frac{u^2+v^2}{2}\} \\ -\frac{1}{2\pi}\frac{v}{u^2+v^2} & \frac{1}{2\pi}\frac{u}{u^2+v^2} \end{vmatrix}$$

$$= -\frac{1}{2\pi}\exp\{-\frac{u^2+v^2}{2}\}.$$

所以, (η_1, η_2)的联合密度函数为

$$\begin{aligned} q(u,v) =& p(x(u,v), y(u,v))|J| \\ =& \frac{1}{2\pi}\exp\{-\frac{u^2+v^2}{2}\}, \quad u, v \in \mathbb{R}. \end{aligned}$$

容易求得: η_1, η_2都服从$N(0,1)$分布, 且相互独立.

*例14 设$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_n)' = \boldsymbol{C}\boldsymbol{\xi} + \boldsymbol{a}$, 其中$\boldsymbol{C}$为$n \times n$可逆矩阵. 求$\boldsymbol{\eta}$的分布.

*例14 设$\boldsymbol{\xi}=(\xi_1,\cdots,\xi_n)'\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$, $\boldsymbol{\eta}=(\eta_1,\cdots,\eta_n)'=\boldsymbol{C}\boldsymbol{\xi}+\boldsymbol{a}$, 其中$\boldsymbol{C}$为$n\times n$可逆矩阵. 求$\boldsymbol{\eta}$的分布.

解: $\boldsymbol{\xi}$的密度函数为

$$p_{\boldsymbol{\xi}}(\boldsymbol{x})=\frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\Big\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big\}.$$

令$\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{a}$, 则$\boldsymbol{x}=\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})$. 所以$\boldsymbol{\eta}$的密度函数为

$$
\begin{aligned}
p_{\boldsymbol{\eta}}(\boldsymbol{y}) =& p_{\boldsymbol{\xi}}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a}))||\boldsymbol{C}^{-1}|| \\
=& \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}||\boldsymbol{C}||} \\
& \cdot \exp\left\{-\frac{1}{2}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a})-\boldsymbol{\mu})\right\} \\
=& \frac{1}{(2\pi)^{n/2}|\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'|^{1/2}} \\
& \cdot \exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{\mu})'(\boldsymbol{C}^{-1})'\boldsymbol{\Sigma}^{-1}\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{\mu})\right\} \\
=& \frac{1}{(2\pi)^{n/2}|\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'|^{1/2}} \\
& \cdot \exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{\mu})'(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{C}\boldsymbol{\mu})\right\}.
\end{aligned}
$$

所以

$$\boldsymbol{\eta} \sim N(\boldsymbol{a}+\boldsymbol{C}\boldsymbol{\mu}, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}').$$

例15 假设随机变量X和Y相互独立, 并且Z是X的函数, W是Y的函数: $Z = g(X)$, $W = h(Y)$. 其中g和h都是Borel函数, 证明: Z和W相互独立.

例15 假设随机变量X和Y相互独立, 并且Z是X的函数, W是Y的函数: $Z = g(X)$, $W = h(Y)$. 其中g和h都是Borel函数, 证明: Z和W相互独立.

证明: 对于任意给定的实数z和w, 定义

$$A = \{x : g(x) \leq z\}, \quad B = \{y : h(y) \leq w\}.$$

A和B是Borel集(因为g和h是Borel函数).

例15 假设随机变量X和Y相互独立, 并且Z是X的函数, W是Y的函数: $Z=g(X)$, $W=h(Y)$. 其中g和h都是Borel函数, 证明: Z和W相互独立.

证明: 对于任意给定的实数z和w, 定义

$$A=\{x: g(x)\le z\},\quad B=\{y: h(y)\le w\}.$$

A和B是Borel集(因为g和h是Borel函数). 又因为X和Y相互独立, 所以

$$\mathsf{P}(X\in A, Y\in B)=\mathsf{P}(X\in A)\mathsf{P}(Y\in B).$$

由此可知对任意的实数$z, w \in \mathbb{R}$,

$$\begin{aligned}\mathsf{P}(Z \le z, W \le w) =& \mathsf{P}(g(X) \le z, h(Y) \le w)\\ =& \mathsf{P}(X \in A, Y \in B)\\ =& \mathsf{P}(X \in A)\mathsf{P}(Y \in B)\\ =& \mathsf{P}(Z \le z)\mathsf{P}(W \le w).\end{aligned}$$

所以W和Z相互独立.

关于随机变量函数的独立性, 我们有更一般的结论:

关于随机变量函数的独立性, 我们有更一般的结论:

定理

令$1 \le n_1 < n_2 < \cdots < n_k = n$; f_1是n_1个变量的Borel函数, f_2是$n_2 - n_1$个变量的Borel函数,$\cdots$, f_k是$n_k - n_{k-1}$个变量的Borel函数. 如果$X_1, X_2, \cdots, X_n$是独立随机变量, 则

$$f_1(X_1, \cdots, X_{n_1}), f_2(X_{n_1+1}, \cdots, X_{n_2}), \cdots, f_k(X_{n_{k-1}+1}, \cdots, X_{n_k})$$

是相互独立的.

2.6.5 数理统计中的几个重要分布

2.6.5 数理统计中的几个重要分布

我们将介绍三个重要的分布: χ^2分布, t分布以及F分布. 这三个分布在数理统计中有很广泛的应用.

2.6.5 数理统计中的几个重要分布

我们将介绍三个重要的分布: χ^2分布, t分布以及F分布. 这三个分布在数理统计中有很广泛的应用.

χ^2分布其实是Γ分布的一个特例, 我们先来讨论Γ分布. $\Gamma(\lambda, r)$ 的密度函数为

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

Γ分布具有一个重要的性质：可加性.

Γ分布具有一个重要的性质：可加性.

引理（Γ分布的可加性）

Γ分布$\Gamma(\lambda, r)$对第二个参数具有可加性：若ξ_1, ξ_2相互独立，且$\xi_1 \sim \Gamma(\lambda, r_1)$，$\xi_2 \sim \Gamma(\lambda, r_2)$，则$\xi_1 + \xi_2 \sim \Gamma(\lambda, r_1 + r_2)$.

证明: 利用卷积公式(2.6.4)来求$\eta=\xi_1+\xi_2$的密度函数:

当$z\le 0$时, $p_\eta(z)=0$.

证明: 利用卷积公式(2.6.4)来求$\eta=\xi_1+\xi_2$的密度函数:

当$z\le 0$时, $p_\eta(z)=0$.

当$z>0$时,

$$
\begin{aligned}
p_\eta(z) &= \int_0^z \frac{\lambda^{r_1}}{\Gamma(r_1)} x^{r_1-1} e^{-\lambda x} \frac{\lambda^{r_2}}{\Gamma(r_2)} (z-x)^{r_2-1} e^{-\lambda(z-x)} \mathrm{d}x \\
&= \frac{\lambda^{r_1+r_2}}{\Gamma(r_1)\Gamma(r_2)} e^{-\lambda z} \int_0^z x^{r_1-1}(z-x)^{r_2-1} \mathrm{d}x \\
&\overset{x=zt}{=} \frac{\lambda^{r_1+r_2}}{\Gamma(r_1)\Gamma(r_2)} e^{-\lambda z} z^{r_1+r_2-1} \int_0^1 t^{r_1-1}(1-t)^{r_2-1} \mathrm{d}t \\
&= \frac{\lambda^{r_1+r_2}}{\Gamma(r_1)\Gamma(r_2)} e^{-\lambda z} z^{r_1+r_2-1} B(r_1,r_2) \\
&= \frac{\lambda^{r_1+r_2}}{\Gamma(r_1+r_2)} e^{-\lambda z} z^{r_1+r_2-1},
\end{aligned}
$$

其中我们用到了贝塔函数的定义和性质:

$$
B(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}dt=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.
$$

综上所述, 可知: $\eta = \xi_1 + \xi_2 \sim \Gamma(\lambda, r_1 + r_2)$.

综上所述, 可知: $\eta = \xi_1 + \xi_2 \sim \Gamma(\lambda, r_1 + r_2)$.

两个特殊的Γ分布:

综上所述, 可知: $\eta=\xi_1+\xi_2\sim\Gamma(\lambda,r_1+r_2)$.

两个特殊的Γ分布:

- $r=1$时, $\Gamma(\lambda,1)$的密度函数为

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

  所以$\Gamma(\lambda,1)$为参数为λ的指数分布.

综上所述, 可知: $\eta=\xi_1+\xi_2\sim\Gamma(\lambda, r_1+r_2)$.

两个特殊的Γ分布:

- $r=1$时, $\Gamma(\lambda,1)$的密度函数为

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

  所以$\Gamma(\lambda,1)$为参数为λ的指数分布.

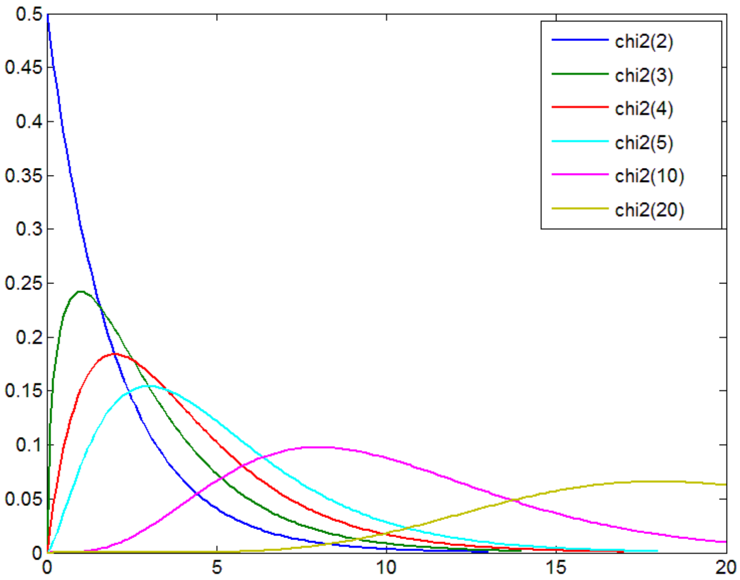
- 当$\lambda=1/2, r=n/2$(n为自然数)时, $\Gamma(1/2,n/2)$是将要介绍的$\chi^2(n)$ 分布.

1. χ^2分布

1. χ^2分布

称$\Gamma(1/2, n/2)$为$\chi^2(n)$分布, 其中称n为它的自由度(degree of freedom/df). 它的密度函数为

$$p(x) = \begin{cases} \frac{(1/2)^{n/2}}{\Gamma(n/2)} x^{n/2-1} e^{-x/2}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$


chi2(2)
chi2(3)
chi2(4)
chi2(5)
chi2(10)
chi2(20)
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20

定理

(1) χ^2分布具有可加性，也就是说，当$\xi_1 \sim \chi^2(n_1)$，$\xi_2 \sim \chi^2(n_2)$，且ξ_1, ξ_2相互独立时，$\xi_1 + \xi_2 \sim \chi^2(n_1 + n_2)$.

(2) 若$\xi_1, \cdots, \xi_n$相互独立，都服从$N(0,1)$，则

$$\eta = \xi_1^2 + \cdots + \xi_n^2 \sim \chi^2(n).$$

定理

(1) χ^2分布具有可加性, 也就是说, 当$\xi_1 \sim \chi^2(n_1)$, $\xi_2 \sim \chi^2(n_2)$, 且ξ_1, ξ_2相互独立时, $\xi_1 + \xi_2 \sim \chi^2(n_1 + n_2)$.

(2) 若$\xi_1, \cdots, \xi_n$相互独立, 都服从$N(0,1)$, 则

$$\eta = \xi_1^2 + \cdots + \xi_n^2 \sim \chi^2(n).$$

证明: (1) 由Γ分布的可加性可推出χ^2分布的可加性.

定理

(1) χ^2分布具有可加性，也就是说，当$\xi_1 \sim \chi^2(n_1)$, $\xi_2 \sim \chi^2(n_2)$，且ξ_1, ξ_2相互独立时，$\xi_1 + \xi_2 \sim \chi^2(n_1 + n_2)$.

(2) 若$\xi_1, \cdots, \xi_n$相互独立，都服从$N(0,1)$，则

$$\eta = \xi_1^2 + \cdots + \xi_n^2 \sim \chi^2(n).$$

证明: (1) 由Γ分布的可加性可推出χ^2分布的可加性.

(2) 记$\eta_i = \xi_i^2, i = 1, \cdots, n$，可知$\eta_i, i = 1, \cdots, n$是相互独立的. 我们曾在例4中得到$\eta_i$的密度函数为

$$p_{\eta_i}(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

因为$\Gamma(1/2)=\sqrt{\pi}$, 所以$\eta_i \sim \chi^2(1)$. 再结合(1)可得

$$\eta=\sum_{i=1}^{n}\xi_i^2=\sum_{i=1}^{n}\eta_i \sim \chi^2(n).$$

因为$\Gamma(1/2)=\sqrt{\pi}$, 所以$\eta_i \sim \chi^2(1)$. 再结合(1)可得

$$\eta=\sum_{i=1}^{n}\xi_i^2=\sum_{i=1}^{n}\eta_i \sim \chi^2(n).$$

注: χ^2分布的自由度n即是$\sum_{i=1}^{n}\xi_i^2$中独立的随机变量的个数.

2. t分布

2. t分布

定理

若$\xi \sim N(0,1)$, $\eta \sim \chi^2(n)$, 且ξ, η相互独立. 则随机变量$T = \frac{\xi}{\sqrt{\eta/n}}$的密度函数为

$$p(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1 + x^2/n)^{-(n+1)/2}, \quad x \in \mathbb{R}.$$

称具有上述密度函数的随机变量T服从$t(n)$分布, 记为$T \sim t(n)$, n为它的自由度.

2. t分布

定理

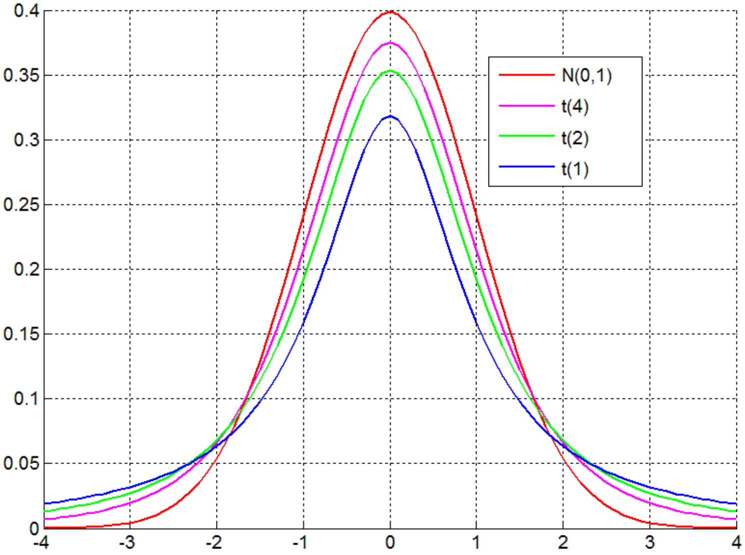
若$\xi \sim N(0,1)$, $\eta \sim \chi^2(n)$, 且ξ,η相互独立. 则随机变量$T = \frac{\xi}{\sqrt{\eta/n}}$的密度函数为

$$p(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2}, \quad x \in \mathbb{R}.$$

称具有上述密度函数的随机变量T服从$t(n)$分布, 记为$T \sim t(n)$, n为它的自由度.

注: 密度函数的推导可参见本章附录. 另外, 利用Γ函数的性质($\Gamma(x) \sim \sqrt{2\pi}e^{-x}x^{x+1/2}$)可以证明

$$\lim_{n\to\infty} \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2} = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \quad \forall x \in \mathbb{R}.$$



N(0,1)
t(4)
t(2)
t(1)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-4
-3
-2
-1
0
1
2
3
4

3. F分布

3. F分布

定理

设随机变量$\xi \sim \chi^2(m)$, $\eta \sim \chi^2(n)$, 且ξ, η相互独立. 则$F = \frac{\xi/m}{\eta/n}$的密度函数为

$$p(x) = \begin{cases} \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)} m^{m/2} n^{n/2} \frac{x^{m/2-1}}{(mx+n)^{(m+n)/2}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

称具有上述密度函数的随机变量F服从$F(m,n)$分布, 记为$F \sim F(m,n)$, m, n分别为它的第一自由度和第二自由度.

3. F分布

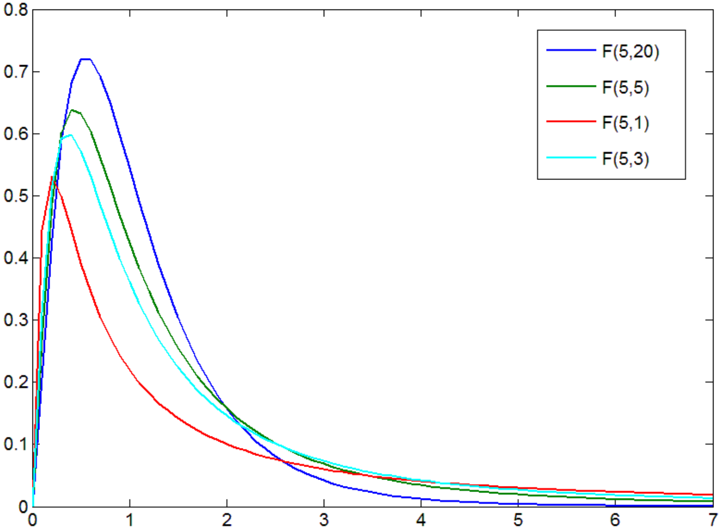
定理

设随机变量$\xi \sim \chi^2(m)$, $\eta \sim \chi^2(n)$, 且ξ, η相互独立. 则$F = \frac{\xi/m}{\eta/n}$的密度函数为

$$p(x) = \begin{cases} \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)} m^{m/2} n^{n/2} \frac{x^{m/2-1}}{(mx+n)^{(m+n)/2}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

称具有上述密度函数的随机变量F服从$F(m,n)$分布, 记为$F \sim F(m,n)$, m, n分别为它的第一自由度和第二自由度.

注: 密度函数的推导可参见本章附录.


F(5,20)
F(5,5)
F(5,1)
F(5,3)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 1 2 3 4 5 6 7

F分布性质:

- 若$F \sim F(m,n)$, 则$\frac{1}{F} \sim F(n,m)$.
- 若$T \sim t(n)$, 则$T^2 \sim F(1,n)$.

F分布性质:

- 若$F \sim F(m, n)$, 则$\frac{1}{F} \sim F(n, m)$.
- 若$T \sim t(n)$, 则$T^2 \sim F(1, n)$.

证明: 第一个结论可由F分布的定义得到. 对于第二个结论, 因为

$$T^2 = \xi^2/(\eta/n),\ \xi^2 \sim \chi^2(1), \eta \sim \chi^2(n),$$

且ξ^2, η相互独立. 所以可知$T^2 \sim F(1, n)$.