概率论

笔记整理

姓名：刘斯宇
学号：17341110

目录

1 概率论的基本概念 **4**
1.1 随机现象 4
1.2 样本空间 4
1.3 事件的关系和运算 5
1.4 随机事件及其概率 7
1.5 条件概率 8
1.6 独立性 10

2 随机变量及其分布 **12**
2.1 离散型随机变量 12
2.2 连续型随机变量 15
2.3 分布函数 17
2.4 复合函数的分布函数 20

3 多维随机变量及其分布 **22**
3.1 两个常见的二维分布 23
3.1.1 均匀分布 23
3.1.2 正态分布 23
3.2 条件分布 24
3.3 随机变量的独立性 27
3.4 已知联合分布，求函数的分布 28
3.4.1 连续型随机变量和的分布 28
3.4.2 连续型随机变量商的分布 30

3.4.3 随机变量的 max 和 min 分布 . . . 31
3.5 二维随机变量的推广 . . . 32

4 随机变量的数字特征 **33**
4.1 一维随机变量的数字特征 . . . 33
4.2 二维随机变量的数字特征 . . . 35
4.3 方差 . . . 36
4.4 协方差与相关系数 . . . 41
4.5 随机变量的其它数字特征 . . . 43

5 大数定理及中心极限定理 **45**
5.1 大数定律 . . . 45
5.2 中心极限定理 . . . 46

6 样本及抽样分布 **49**
6.1 卡方分布 . . . 51
6.2 t 分布 . . . 53
6.3 F 分布 . . . 54

7 参数估计 **57**

第一章--概率论的基本概念

1 概率论的基本概念

1.1 随机现象

基本概念

- 在一定条件下可能出现也可能不出现的现象称为**随机现象**
- **随机试验**: 若试验满足
  - 可在相同的条件下重复进行;
  - 试验的可能结果不止一个, 但事先能明确所有可能发生的结果;
  - 试验前不能预知出现哪种结果

  称这样的试验为随机试验。
- **随机事件**：随机试验的结果，称为随机事件，简称为事件。用大写字母 A，B，C 等表示.
- 一定条件下必然发生的事件称为**必然事件**，用 Ω 表示
- 一定条件下必然不发生的事件称为**不可能事件**，用 Φ 表示
- **样本点**: 试验的每一个可能发生的结果称为一个样本点, 记为 ω.
- 样本空间：随机试验的所有可能结果组成的集合称为样本空间，记为 Ω

1.2 样本空间

基本概念

- **样本点**: 试验的每一个可能发生的结果称为一个样本点, 记为 ω.
- 样本空间：随机试验的所有可能结果组成的集合称为样本空间，记为 Ω
-

$$\text{样本空间}\begin{cases}\text{有限样本空间}\\ \text{无限样本空间}\begin{cases}\text{可数样本空间} & \text{不停地数，总有一天能数完所有的样本}\\ \text{不可数样本空间} & \text{不停地数，总有一个样本数不到}\end{cases}\end{cases}$$

Tip 可数样本空间的例子：正整数，不可数样本空间的例子：0-1 之间的实数

1.3 事件的关系和运算

基本概念

- 包含关系: 如果事件 A 发生必然导致事件 B 发生，则称事件 B 包含事件 A，记作 $A \subset B$
- 相等关系: 若 $A \subset B$,$B \subset A$，则称事件 A 和事件 B 相等，记作 $A = B$
- 事件的和：事件 A、B 至少有一个发生构成的事件称为事件 A 和事件 B 的和，记作 $A \cup B$
- 事件的积：两事件 A、B 同时发生也是一个事件，称为事件 A 和事件 B 的积事件，记作 $A \cap B$, 也记作 AB
- 事件的余：设 A 是事件，称“非 A”是 A 的余事件（或对立事件）. 含义是：“非 A”发生当且仅当 A 不发生。常用 $\bar{A}$ 表示“非 A”
- 事件的差: 事件“A 出现而 B 不出现”，称为事件 A 与 B 的差. 记作 A- B.
- 事件的互不相容 (互斥)：若事件 A 、B 满足 $A \cap B = AB = V$ 则称事件 A 与 B 互不相容，其中 V 表示不可能事件.

Properties: 事件的关系和运算

- 若两个事件是对立事件，那么他们一定互斥，但是如果两个事件是互斥，则他们不一定对立。**因为对立还要求 A+B=U, 但是互斥并没有这个要求**
- 交换律
$$A \cup B = B \cup A, AB = BA$$
- 结合律
$$(A \cup B) \cup C = A \cup (B \cup C), (AB)C = A(BC)$$
- 分配律
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) = AB \cup AC$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C) = (A \cup C)(B \cup C)$$
- 对欧律
$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}$$

例题 1

证明:

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

解答

$$\begin{aligned}
&\text{Assume } x \in (A \cup B)^c \\
&\Rightarrow x \notin A \cup B \\
&\Rightarrow x \notin A \text{ and } x \notin B \\
&\Rightarrow x \in A^c \text{ and } x \in B^c \Rightarrow x \in A^c \cap B^c \\
&\Rightarrow \quad (A \cup B)^c \subseteq (A^c \cap B^c)
\end{aligned}$$

$$\begin{aligned}
&\text{Assume } x \in A^c \cap B^c \\
&\Rightarrow x \notin A \text{ and } x \notin B \\
&\text{If } x \notin (A \cup B)^c \\
&\Rightarrow x \in A \cup B \Rightarrow x \in A \text{ or } x \in B \\
&\text{Thus } x \in (A \cup B)^c \Rightarrow [(A^c \cap B^c) \subseteq (A \cup B)^c
\end{aligned}$$

推广：多个时间的和与积

- 称 $\cup_{k=1}^{n} A_k$ 为 n 个事件 $A_1, A_2, \ldots, A_n$ 的和事件，即 $A_1, A_2, \ldots, A_n$ 至少发生一个
- 称 $\cup_{k=1}^{\infty} A_k$ 为可列个事件 $A_1, A_2, \ldots$ 的和事件，即 $A_1, A_2, \ldots$ 至少发生一个
- 称 $\cap_{k=1}^{n} A_k$ 为 n 个事件 $A_1, A_2, \ldots, A_n$ 的积事件，即 $A_1, A_2, \ldots, A_n$ 同时发生
- 称 $\cap_{k=1}^{\infty} A_k$ 为可列个事件 $A_1, A_2, \ldots$ 的积事件，即 $A_1, A_2, \ldots$ 同时发生

1.4 随机事件及其概率

基本概念

- 发生的频率有稳定性的事件称为随机事件，简称事件，频率摆动的中心叫作该随机事件的概率。
- 古典概率: 具有下面特点的试验称为古典概率。
  - 试验的样本空间只包含有限个样本点
  - 试验中每个基本时间发生的可能性相同

$$P(A) = \frac{N(A)}{N(S)}$$

- 几何概率: 有无穷多个等可能结果的随机事件
- 概率的公理化定义与性质
  - 设 S 是样本空间，事件 a 是样本空间中的一个样本，定义一个实数 $P = P(A)$，若 P 的定义域和函数 P 满足下列条件
    - 非负性，$P(A) \geq 0$
    - 规范性，$P(S) = 1$
    - 若 $A_n(n \geq 1)$ 且两两互不相容，则

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P\left(A_n\right)$$

Properties: 概率的一些性质

- $P(\Phi) = 0$
- 设有限个事件 $A_1, A_2, \ldots, A_n$ 两两互斥。则

$$P(A_1 + A_2 + \cdots + A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

- $P(\bar{A}) = 1 - P(A)$
- 设 A、B 是两个事件，且 $A \subset B$，则 $P(B - A) = P(B) - P(A)$
- $P(B - A) = P(B) - P(AB)$
- 加法公式

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

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

例题 2

事件 A，B 的概率分别为 $\frac{1}{3}, \frac{1}{2}$，求在下列情况下的 $P(B\bar{A})$ 的值

- (1)$A \subset B$
- (2)$P(AB) = \frac{1}{8}$

解答

1.5 条件概率

基本概念

- 设 A、B 为两个事件，P(B)>0，则称 $\frac{P(AB)}{P(B)}$ 为在事件 B 发生的条件下事件 A 发生的概率，记作

$$P(A|B) = \frac{P(AB)}{P(B)}$$

Properties: 条件概率的一些性质

- 有界性 $0 \leq P(A|B) \leq 1$
- 规范性 $P(B|B) = 1, P(\Phi|B) = 0$
- 可加可列性：设 $A_1, A_2, \ldots$ 是两两不相容的事件，则有

$$P\left(\bigcup_{i=1}^{\infty} A_i|B\right) = \sum_{i=1}^{\infty} P\left(A_i|B\right)$$

定理 1.1 (乘法定理) 设 A、B 是两个事件，若 $P(B) > 0$ 则

$$P(AB) = P(B)P(A|B) = P(A)P(B|A)$$

进一步的，

$$P\left(A_1 A_2 \cdots A_n\right) = P\left(A_1\right) P\left(A_2|A_1\right) \cdots P\left(A_n|A_1 A_2 \cdots A_{n-1}\right)$$

例题 3

场精彩的足球赛将要举行, 5 个球迷好不容易才搞到一张入场券. 大家都想去, 只好用抽签的方法来解决. 后抽比先抽的确实吃亏吗?

解答

我们用 A_i 表示“第 i 个人抽到入场券”,$i = 1,2,3,4,5$. 则 $P(A_1) = \frac{1}{5}, P(\bar{A}_1) = \frac{4}{5}$;

由于 $A_2 = \bar{A}_1 A_2$, 那么由乘法公式我们可以得到 $P(A_2) = P(\bar{A}_1)P(A_2|\bar{A}_1) = \frac{4}{5} * \frac{1}{4} = \frac{1}{5}$

同理可得 $P(A_3) = P(A_4) = P(A_5) = \frac{1}{5}$

定理 1.2 (全概公式) 如果事件组 $B_1, B_2, \ldots, B_n$ 满足

- (1)$B_1, B_2, \ldots, B_n$ 两两互不相容
- (2)$P(B_i) > 0 (i = 1, 2, \ldots)$
- (3)$\cup_{i=1}^{n} B_i = S$(S 表示必然事件) 则对任意事件 A，有

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

证明

$A = AS = A(B_1 \cup)$

例题 4

、乙、丙三人同时对飞机进行射击，三人击中的概率分别为 0.4、0.5、0.7. 飞机被一人击中而击落的概率为 0.2, 被两人击中而击落的概率为 0.6, 若三人都击中, 飞机必定被击落, 求飞机被击落的概率.

解答

定理 1.3 (贝叶斯公式) 如果事件组 $B_1, B_2, \ldots, B_n$ 满足

- (1)$B_1, B_2, \ldots, B_n$ 两两互不相容
- (2)$P(B_i) > 0 (i = 1, 2, \ldots)$
- (3)$\cup_{i=1}^{n} B_i = S$(S 表示必然事件) 则对任意事件 A，有

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

1.6 独立性

基本概念

设 A、B 为两事件，若 $P(AB) = P(A)P(B)$，则称事件 A 和事件 B 独立

Tip 三事件 A, B, C 相互独立是指下面的关系式同时成立：
$$\begin{cases} P(AB) = P(A)P(B) \\ P(AC) = P(A)P(C) \\ P(BC) = P(B)P(C) \\ P(ABC) = P(A)P(B)P(C) \end{cases}$$

基本概念

n 个事件 $A_1, A_2, \ldots, A_n$ 相互独立是指下面的关系式同时成立

$$
\begin{aligned}
&P\left(A_i A_j\right) = P\left(A_i\right) P\left(A_j\right), 1 \leq i < j \leq n \\
&P\left(A_i A_j A_k\right) = P\left(A_i\right) P\left(A_j\right) P\left(A_k\right), 1 \leq i < j < k \leq n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&P\left(A_1 A_2 \cdots A_n\right) = P\left(A_1\right) P\left(A_2\right) \cdots P\left(A_n\right)
\end{aligned}
$$

2 随机变量及其分布

2.1 离散型随机变量

定义 2.1 如果所有可能的结果组成的集合 $S=\{e\}$，$X=X(e)$ 是 S 上有定义的实值函数，而且对任何实数 c, 事件 e: $X(e)<=c$ 是有概率的，则称 X 为随机变量。随机变量通常用大写字母 X,Y,Z,W,N 等表示, 而表示随机变量所取的值时, 一般采用小写字母 x, y, z, w, n 等.(其实随机变量就是一种实值的映射)

定义 2.2 (离散型随机变量) 某些随机变量 X 的所有可能取值是有限多个或可列无限多个, 这种随机变量称为离散型随机变量。

定义 2.3 (离散型随机变量的概率分布) 随机变量 X 所取的一切可能值为 x_k $(k=1,2,\ldots)$，则称 $P\{X=x_k\}\ p_k(k=1,2,\ldots)$ 为离散型随机变量 X 的概率分布，其中 $P_k(k=1,2\ldots)$ 满足

- $p_k \geq 0, k=1,2\ldots$
- $\sum_k p_k = 1$

例题 5

如果随机变量 X 的分布律为 $P\{X=k\}=c\frac{\lambda^k}{k!}(k=1,2\ldots,\lambda>0)$ 试确定未知常数 c。

解答

由分布律的性质可知 $\sum_{k=1}^{\infty}c\frac{\lambda^k}{k!}=\sum_{k=0}^{\infty}c\frac{\lambda^k}{k!}-1=e^{\lambda}-1=1$

二项分布中最可能出现次数的定义与推导

$$\left.\begin{array}{l}\frac{p_{k-1}}{p_k}=\frac{(1-p)k}{p(n-k-1)}\leq 1\\ \frac{p_k}{p_{k+1}}=\frac{(1-p)(k+1)}{p(n-k)}\geq 1\end{array}\right\}\rightarrow (n+1)p-1\leq k\leq (n+1)p$$

定理 2.1 (泊松定理) 设随机变量 X 服从二项分布，又设 $np=\lambda$，则有

$$\begin{aligned}\lim_{n\to\infty}P\{X=k\}&=\lim_{n\to\infty}\binom{n}{k}p^k(1-p)^{n-k}\\&=\frac{\lambda^k e^{-\lambda}}{k!},k=0,1,2,\ldots,n\end{aligned}$$

表 1: 汇总

名称	简写	表达式	运用范围
两点分布 (伯努利分布，0-1 分布)		$P(X)=\begin{cases} p & X=1 \\ 1-p & X=0 \end{cases}$	凡是随机试验只有两个可能的结果，常用 0 –1 分布描述，如产品是否格、人口性别统
二项分布	B(n,p)	$P(X=k)=C_n^k p^k(1-p)^{n-k}$ $k=0,1,\cdots,n$	
泊松分布	$\pi(\lambda)$	$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$, $k=0,1,2,\cdots$	泊松分布适合于描述单位时间（或空间）内随机事件发生的次数. 如某一服务设施在一定时间内到达的人数，电话交换机接到呼叫的次数，汽车站台的候客人数，机器出现的故障数
超几何分布		$P(X=k)=\frac{C_D^k C_{N-D}^{n-k}}{C_N^n}$ $(k=0,1,\ldots,n)$	超几何分布的极限分布就是二项分布. 用于数量不是无限的。
几何分布		$P(X=k)=p(1-p)^{k-1}$ $(k=1,\ldots,n)$	无记忆性。

证明泊松定理

$$\begin{aligned}
\lim_{n\to\infty} P(X=k) &= \lim_{n\to\infty}\begin{pmatrix} n \\ k \end{pmatrix} p^k(1-p)^{n-k} \\
&= \lim_{n\to\infty}\frac{n!}{(n-k)!k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \lim_{n\to\infty}\left[\frac{n!}{n^k(n-k)!}\right]\left(\frac{\lambda^k}{k!}\right)\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k} \\
&= \lim_{n\to\infty}\left[\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right]\left(\frac{\lambda^k}{k!}\right)\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k} \\
&= \left(\frac{\lambda^k}{k!}\right)\exp(-\lambda)
\end{aligned}$$

例题 6

设同类型设备 90 台, 每台工作相互独立，每台设备发生故障的概率都是 0.01. 在通常情况下，一台设备发生故障可由一个人独立维修, 每人同时也只能维修一台设备.

- (1) 问至少要配备多少维修工人，才能保证当设备发生故障时不能及时维修的概率小于 0.01?
- (2) 问 3 个人各自独立负责 30 台设备，设备发生故障不能得到及时维修的概率?

- (3) 问 3 个人共同负责 90 台，设备发生故障不能得到及时维修的概率?

解答

- (1) 设需要配备 N 个维修工人, 设 X 为 90 台设备中发生故障的台数，则 X ~B(90, 0.01).$P(X>N)=\sum_{k=N+1}^{90}C_{90}^{k}(0.01)^{k}(0.99)^{N-k}$, 令 $\lambda=90*0.01=0.9$，则

$$\begin{aligned}P(X>N)&\approx\sum_{k=N+1}^{90}e^{-0.9}\frac{0.9^{k}}{k!}\\&=\sum_{k=N+1}^{\infty}e^{-0.9}\frac{0.9^{k}}{k!}-\sum_{k=9}^{\infty}e^{-0.9}\frac{0.9^{k}}{k!}\\&\approx\sum_{k=N+1}^{\infty}e^{-0.9}\frac{0.9^{k}}{k!}<0.01\end{aligned}$$

  查表可以知道，N=4

- (2) 设 30 台设备中发生故障的台数为 Y，则 Y ~B(30, 0.01). 设每个人独立负责 30 台设备, 第 i 个人负责的 30 台设备发生故障不能及时维修为事件 $A_i(i=1,2,3)$，则 $P(A_i)=P(Y\geq 2)\approx\sum_{k=2}^{\infty}e^{-0,3}\frac{0.3^{k}}{k!}=0.0369$ 三个人各独立负责 30 台设备发生故障不能及时维修为事件 $A_1\cup A_2\cup A_3$

$$P(A_1\cup A_2\cup A_3)=1-P(\overline{A_1\cup A_2\cup A_3})=1-\prod_{i=1}^{3}P(\bar{A}_i)$$

$$=1-(1-0.0369)^{3}\approx 0.1067$$

- (3) 设 X 为 90 台设备中发生故障的台数，则 X ~B(90, 0.01). 三个人共同负责 90 台设备发生故障不能及时维修的概率为

$$\begin{aligned}P(X>3)&\approx\sum_{k=4}^{90}e^{-0.9}\frac{0.9^{k}}{k!}\\&=\sum_{k=4}^{\infty}e^{-0.9}\frac{0.9^{k}}{k!}-\sum_{k=91}^{\infty}e^{-0.9}\frac{0.9^{k}}{k!}\\&\approx\sum_{k=4}^{\infty}e^{-0.9}\frac{0.9^{k}}{k!}\\&=0.013459<0.1067\end{aligned}$$

  故三个人共同负责 90 台设备比各自负责好得多!

2.2 连续型随机变量

定义 2.4 (连续型随机变量) 设 X 是随机变量，如果存在非负函数 $p(x)$，满足：对于任意的 a,b，有 $P\{a<X<b\}=\int_a^b p(x)dx$，则称 X 是连续型随机变量，称为 X 的概率密度函数 *probability density function*（小写 *pdf*），简称概率密度.

推论 2.1 对于任意可能值 a，连续型随机变量取 a 的概率等于零. 即 $p\{X=a\}=0$, 因此 $P\{a\le X\le b\}=P\{a<X\le b\}=P\{a\le X<b\}=P\{a<X<b\}$

证明

$$P(X=a)=\lim_{\substack{\Delta x\to 0\\ a+\Delta x}} P(a-\Delta x<X<a+\Delta x)$$
$$=\lim_{\Delta x\to 0}\int_{a-\Delta x}^{a+\Delta x} p(x)dx=\mathbf{0}$$

Tip p (a) 不是 X =a 的概率！密度函数 p (x) 在某点处 a 的高度，并不反映 X 取值的概率. 但是，这个高度越大，则 X 取 a 附近的值的可能性就越大. 也可以说，在某点密度曲线的高度反映了概率集中在该点附近的程度.

Tip 概率为零的事件不一定是不可能事件。根据上面的讨论 $p\{X=a\}=0$, 但是 $\{X=a\}$ 并非不可能事件，也就是说 $P(A)=0 \nRightarrow A=\Phi$ 和 $P(B)=1\nRightarrow B=\Omega$

证明指数分布的无记忆性

即证明：$P(T>s+t\mid T>t)=P(T>s)$ for all $s,t\ge 0$.

$$\boldsymbol{P\{X>s+t\mid X>s\}=\frac{P\{(X>s+t)\cap(X>s)\}}{P\{X>S\}}}$$

$$=\frac{\boldsymbol{P\{X>s+t)\}}}{\boldsymbol{P\{X>S\}}}=\frac{\int_{s+t}^{\infty}\lambda e^{-\lambda x}dx}{\int_{s}^{\infty}\lambda e^{-\lambda x}dx}=\frac{-e^{-\lambda x}\big|_{s+t}^{\infty}}{-e^{-\lambda x}\big|_{s}^{\infty}}$$

$$=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=e^{-\lambda t}$$

表 2: 连续型随机变量的分布汇总

名称	简写	表达式	意义
均匀分布	U(a,b)	$p(x)=\begin{cases}\frac{1}{b-a}, & a\leq x\leq b\\ 0, & others\end{cases}$	在区间 (a,b) 上服从均匀分布的随机变量 X，落在区间 (a,b) 中任意长度的子区间内的可能性是相同的。例如对小数点后第一位进行四舍五入时，那么一般认为误差服从（-0.5, 0.5）上的均匀分布。
指数分布	U(a,b)	$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\leq 0\end{cases}$	无记忆性，有些系统的寿命分布也可用指数分布来近似, 当电子产品的失效是偶然失效时, 其寿命服从指数分布. 在排队论中它被广泛地用于描绘等待时间, 如电话通话时间、各种随机服务系统的服务时间、等待时间等.
高斯分布	$N(\mu,\sigma^2)$	$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x\in R$	

指数分布和几何分布的无记忆性怎么理解？

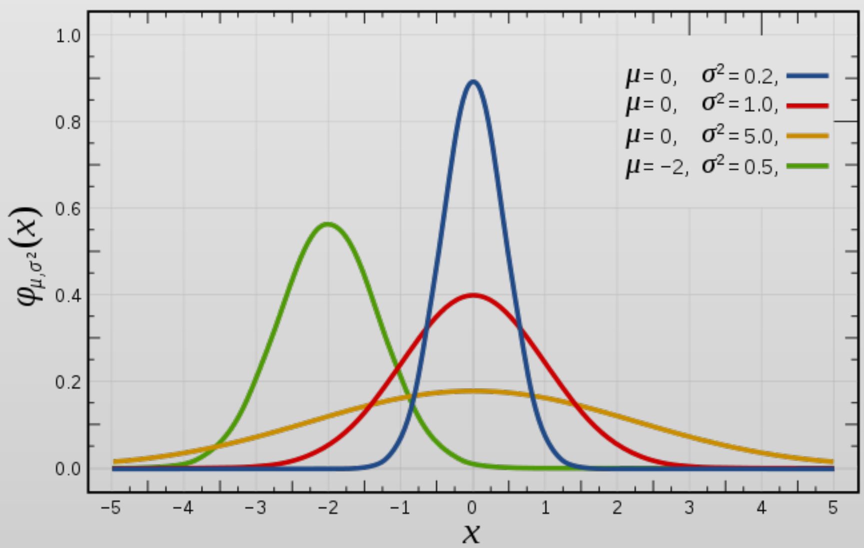
举个例子，比如我在等公交车，如果等公交车的时间服从指数分布，那么给定我已经等了 5 分钟，我需要继续等，到第 11 分钟公交车还没来的概率，与等 6 分钟公交车没有来的概率相等：未来我还需要等多长时间，跟我已经等了多长时间没有关系。

无记忆性的离散版本是几何分布。比如北京的车牌，假设每次车牌抽签抽中的概率都相等，那么一个一次都没有抽签抽过的人，跟一个抽了 50 次的人，下一次抽签抽中的概率都是一样的，从而不管一个人抽签抽过多少次，未来还需要抽的次数的分布是一样的，这就是无记忆性。

Tip 无记忆分布一定是指数分布 (exponential distribution) 或几何分布 (geometric distribution); 其他分布都没有无记忆性.因为满足 $f(a+b)=f(a)f(b)$ 的函数只有指数函数

高斯分布的性质

- 图形关于直线 $x = \mu$ 对称, 即 $f(\mu + x) = f(\mu - x)$
- 在 $x = \mu + \sigma$ 时, 曲线 y = f (x) 在对应的点处有拐点（即曲线的凹凸分界点）.
- μ —位置参数, 固定 σ , 对于不同的 μ, 对应的 f (x) 的形状不变化，只是位置不同。
- σ—形状参数, 固定 μ , 对于不同的 σ, 对应 f (x) 位置不变化, 只是 σ 越小，靠近 μ 附近取值的概率越大，相应拐点越接近于 ox 轴.



定理 2.2 (正态分布的 3σ 定理) $|X - \mu| > 3\sigma$ 的概率很小，因此可认为正态随机变量的取值几乎全部集中在 $[\mu - 3\sigma, \mu + 3\sigma]$ 区间内

2.3 分布函数

定义 2.5 (离散型变量的分布函数) 设 X 是随机变量，称 $F(x) = P(X \leq x)$ 为 X 的分布函数。

> **Tip** 注意分布函数的定义里面取等了。

分布函数的性质

- (非负性)$0 \leq F(x) \leq 1$
- (单调性)F(x) 是 x 的非减函数
- (规范性)$F(-\infty) = 0, F(\infty) = 1$
- (右连续性)$\lim_{x \to x_0^+} F(x) = F(x_0)$ 也就是说 $F(X + 0^+) = F(X)$

分布函数的右连续性怎么理解？

首先随机变量的分布函数是允许有间断点的，此时代表随机变量落在那个点上的概率（不是概率密度）不等于 0。对这种点，如果规定分布函数的函数值等于右极限，则它右连续；如果规定分布函数的函数值等于左极限，则它左连续。两种规定方法并无优劣差别，通行做法似乎采用前者。之所以分布函数满足右连续性就是因为我们分布函数的定义取等了。

这样说很清楚了，那么我们再来举个例子好了，对于离散型的变量 X，$P(x=0) = 0.5, P(x=1) = 0.5$ 那么它的分布函数为

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.5, & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

定义 2.6 (连续型变量的分布函数) 设 *F(x)* 是随机变量 *X* 的分布函数, 如果存在一非负函数 *p(x)*，使对任意实数 *x* , 有

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(t)dt$$

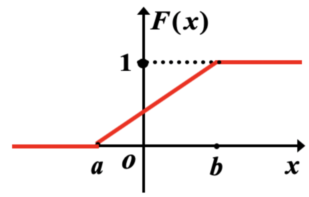
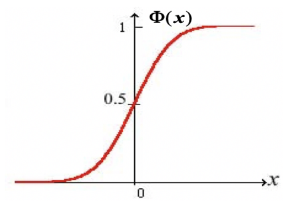
标准正态分布的分布函数的性质

- $\Phi(0) = \frac{1}{2}$
- $\forall x \in R, \Phi(-x) = 1 - \Phi(x)$

$$\begin{aligned} \Phi(-x) &= \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-\frac{t^2}{2}} dt \\ &\overset{u=-t}{=} \tfrac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-\frac{u^2}{2}} du \\ &= 1 - \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du = 1 - \Phi(x) \end{aligned}$$

定理 2.3 若 $X \sim N(\mu, \sigma^2)$, 则 $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

表 3: 随机变量的分布函数汇总

名称	分布函数	图像
均匀分布	$F(x)=P(X\leq x)=\begin{cases}0, & x\leq a\\ \frac{x-a}{b-a}, & a<x\leq b\\ 1, & x>b\end{cases}$	
正态分布	$F(x)=\Phi(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt, x\in R$	

Tip 该定理的重要性在于，任何一个一般的正态分布都可以通过线性变换转化为标准正态分布.

证明

$$
\begin{aligned}
P\{Z\leq x\} &= P\left\{\frac{X-\mu}{\sigma}\leq x\right\}=P\{X\leq\mu+\alpha x\}\\
&=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\mu+\alpha x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt\\
&\overset{u=\frac{t-\mu}{\sigma}}{=}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}d=\Phi(x)
\end{aligned}
$$

推论 2.2 若 $X\sim N(\mu,\sigma^2)$, 则

$$
\begin{aligned}
F(x) =& P\{X\leq x\}=P\left\{\frac{X-\mu}{\sigma}\leq\frac{x-\mu}{\sigma}\right\}\\
&=\Phi\left(\frac{x-\mu}{\sigma}\right)
\end{aligned}
$$

例题 7

公共汽车车门的高度是按男子与车门顶头碰头机会在 0.01 以下来设计的. 设男子身高 $X\sim N(170,62)$, 问车门高度应如何确定?

解答

设车门高度为 h cm, 按设计要求 $P(X \geq h) \leq 0.01$, 我们要求的是满足 $P(x < h) \geq 0.99$ 的最小的 h，由于

$$\begin{aligned} P(X < h) &= P\left(\frac{X-170}{6} < \frac{h-170}{6}\right) \\ &= \Phi\left(\frac{h-170}{6}\right) \end{aligned}$$

查表可以得到 $\Phi(2.33) = 0.9901 > 0.99$，所以 $\frac{h-170}{6} = 2.33 \Rightarrow h = 184$

2.4 复合函数的分布函数

定理 2.4 设随机变量 X 具有概率密度 $f_X(x), x \in R$, 又设函数 $g(x)$ 处处可导，且有 $g'(x) > 0$ 或 $g'(x) < 0$ 恒成立，则 $Y = g(X)$ 是一个连续型随机变量 Y，其概率密度为

$$f_Y(y) = \begin{cases} f_X[h(y)]\,|h'(y)|, & \alpha < y < \beta \\ 0, & others \end{cases}$$

其中 $h(y)$ 是 $g(x)$ 的反函数,$\alpha = min\{g(-\infty), g(\infty)\}$,$\beta = max\{g(-\infty), g(\infty)\}$

证明

设随机变量 $Y = g(X)$ 的分布函数为 $F_Y(y)$，则有 $F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\}$, 设 $g'(x) > 0$，则 g(x) 单调递增，因此 $F_Y(y) = P\{X \leq g^{-1}(y)\} = P\{X \leq h(y)\} = \int_{-\infty}^{h(y)} f_X(x)dx$, 运用变限的定积分的求导公式，两边同时求导就可以得到。

推论 2.3 若 $g(x)$ 在不相叠的区间 $I_1, I_2 \ldots$ 上逐段严格单调，其反函数分别为 $h_1(y), h_2(y), \ldots$ 均为连续函数，$Y = g(x)$ 是一个连续型随机变量 Y，其概率密度为

$$f_Y(y) = f_X(h_1(y))\left|h_1'(y)\right| + f_X(h_2(y))\left|h_2'(y)\right| + \cdots$$

推论 2.4 若 $X \sim N(\mu, \sigma^2)$, 则 X 的线性函数 $Y = aX + b$ 也服从正态分布，且 $Y \sim N(a\mu + b, (a\sigma)^2)$

证明

$$\begin{aligned} f_Y(y) &= f_X[h(y)]\,|h'(y)| = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) \\ &= \frac{1}{|a|}\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{a}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma|a|} e^{-\frac{[y-(a\mu+b)^2}{2(a\sigma)^2}} \end{aligned}$$

Tip 记不住公式没关系，其实我们只需要先求出分布函数然后再求导就可以了，这样的方法是万无一失的。

例题 8

设 $V = A\sin\theta$, 其中 A 是一个已知的常数，θ 是一个随机变量，在区间 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 上服从均匀分布，求 V 的概率密度。

解答

$$f(\theta) = \begin{cases} \frac{1}{\pi}, & -\frac{\pi}{2} < \theta < \frac{\pi}{2} \\ 0, & others \end{cases}$$

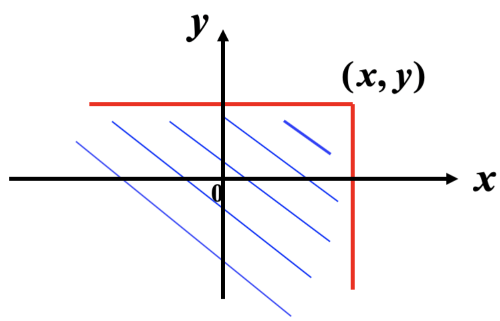
由于 $V = g(\theta) = A\sin\theta$ 在区间 $(-\frac{\pi}{2}, \frac{\pi}{2})$,$g'(\theta)$ 处处存在，且 $g'(\theta) > 0$ 且存在反函数 $h(\theta) = \arcsin\frac{v}{A}$,$h'(v) = \frac{1}{\sqrt{A^2-v^2}}$,$v \in [-A, A]$, 所以

$$f_Y(y) = \begin{cases} \frac{1}{\pi} \times \frac{1}{\sqrt{A^2-v^2}}, & -A < v < A \\ 0, & others \end{cases}$$

3 多维随机变量及其分布

定义 3.1 如果二维随机变量 (X,Y) 全部可能取到的不相同的值是有限对或可列无限多对，则 (X,Y) 是离散型随机变量。记 $P(X=x_i,Y=y_j)=p_{ij}$ 为二维离散型随机变量 (X,Y) 的概率分布，或随机变量 X 和 Y 的联合分布。

定义 3.2 (二维随机变量的分布函数) 设 (X,Y) 是二维随机变量，对于任意的实数 x，y，二元函数 $F(x,y)=P(X\le x,Y\le y)$ 称为二维随机变量 (X,Y) 的分布函数或称为 (X,Y) 的联合分布函数。



二维随机变量分布函数的性质

- F(x,y) 分别对 x 和 y 单调非减
- F(x,y) 对每个自变量 x 或 y 是右连续的
- $0\le F(x,y)\le 1$
- 当 $x_1\le x_2, y_1\le y_2$ 时，有 $F(x_2,y_2)-F(x_2,y_1)+F(x_1,y_1)-F(x_!,y_2)\ge 0$

定义 3.3 (二维连续型随机变量) 设 (X,Y) 是二维随机变量，如果存在定义在平面上的函数 $p(x,y)$，满足条件

- $p(x,y)\ge 0$
- $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p(x,y)dxdy=1$
- G 是 xoy 平面的一块区域，点 (X,Y) 落在 G 内的概率为 $P\{(X,Y)\in G\}=\iint_G p(x,y)\mathrm{d}x\mathrm{d}y$

则称 (X,Y) 是连续型随机变量，而 $p(x,y)$ 称为二维随机变量 (X,Y) 的概率密度函数或称为随机变量 X 和 Y 的联合概率密度函数. 且 $p_X(x)=\int_{-\infty}^{\infty}p(x,y)dy$ 称为 (X,Y) 关于 X 的边缘概率密度，$p_Y(y)=\int_{-\infty}^{\infty}p(x,y)dx$ 称为 (X,Y) 关于 Y 的边缘概率密度。

3.1 两个常见的二维分布

3.1.1 均匀分布

定义 3.4 设 G 是平面上的有界区域, 其面积为 S. 若二维随机变量 (X,Y) 的概率密度为

$$p(x,y)=\begin{cases}\frac{1}{S} & (x,y)\in G\\ 0, & others\end{cases}$$

则称 (X,Y) 在 G 上服从均匀分布.

> **Tip** 均匀分布的边缘密度不再是一维均匀分布。

例子

设随机向量 (X,Y) 服从区域 D 上的均匀分布, 其中 D=(x,y),$x^2+y^2\leq 1$, 但是 X,Y 的边缘密度函数 $p_1(x)$ 和 $p_2(y)$ 为

$$p_1(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2} & ,|x|\leq 1\\ 0 & ,\quad |x|>1\end{cases}$$

$$p_2(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2} & ,|y|\leq 1\\ 0 & ,\quad |y|>1\end{cases}$$

3.1.2 正态分布

定义 3.5 若二维随机变量 (X,Y) 具有概率密度

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2\left(1-\rho^2\right)}\left[\frac{\left(x-\mu_X\right)^2}{\sigma_X^2}+\frac{\left(y-\mu_Y\right)^2}{\sigma_Y^2}-\frac{2\rho\left(x-\mu_X\right)\left(y-\mu_Y\right)}{\sigma_X\sigma_Y}\right]\right)$$

其中 $\mu_X,\mu_Y,\sigma_X>0,\sigma_Y>0,\rho$ 均为常数, 且 $|\rho|<1$, 则称 (X,Y) 服从二维正态分布, 记作 $(X,Y)\sim N(\mu_X,\sigma_X^2,\mu_Y,\sigma_Y^2)$

推论 3.1 二维正态分布的两个边缘分布都是一维正态分布, 并且不依赖于参数 ρ

证明

由于

$$\begin{aligned}&\frac{(y-\mu_2)^2}{\sigma_2^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\\&=\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2-\rho^2\frac{(x-\mu_1)^2}{\sigma_1^1}\end{aligned}$$

那么

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

令 $t=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)$

$$\begin{aligned}p_X(x)=&\frac{1}{2\pi\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=\frac{1}{2\pi\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\cdot\sqrt{2\pi}\\=&\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\quad(-\infty<x<\infty)\end{aligned}$$

同理可得 $p_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}\quad(-\infty<y<\infty)$

3.2 条件分布

定义 3.6 (离散条件分布) 对于离散型的随机变量 X 和 Y (取值范围分别是 $\mathcal{I}$ 和 $\mathcal{J}$)，随机变量 Y 在条件 X =x 下的条件概率分布是：

$$\forall j\in\mathcal{J},\quad p_{Y|X}(j)=p_Y(j\mid X=i)=P(Y=j\mid X=i)=\frac{P(X=i,Y=j)}{P(X=i)}.P(X=i)>0P(X=i)>0$$

定义 3.7 (连续条件分布) 对于连续型的随机变量 X 和 Y，$P(X=i)=P(Y=j)=0$，因此对离散型随机变量的条件分布定义不适用。假设其联合密度函数为 $f(x,y)$，X 和 Y 的边际密度函数分别是 $f_X(x)$ 和 $f_Y(y)$，那么 Y 在条件 X =x 下的条件概率密度函数是：

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

例题 9

一射手进行射击, 每次射击击中目标的概率均为 p (0<p<1) 且假设各次击中目标与否相互独立, 射击进行到击中目标两次为止. 设以 X 表示到第一次击中目标所需要的射击次数, 以 Y 表示总共进行的射击次数. 试求 (X , Y) 的联合分布律和条件分布律.

解答

由题意 {X=i} 表示第 i 次首次击中目标，{Y=j} 表示第 j 次击中目标. 因而 i= j, {X=i, Y=j} 是不可能事件. 即

$$P\{X=i,Y=j\}=0$$

当 i<j，X=i, Y=j 表示第 i 次和第 j 次击中目标而其余 j-2 次均未击中目标. 于是 (X, Y) 的联合分布律为:

$$P\{X=i,Y=j\}=p^2q^{j-2}$$

随机变量 X 的边缘分布律

$$P(X=i)=pq^{i-1}(i=1,2\dots)$$

于是，在 X=i (i=1，2，...) 的条件下，随机变量 Y 的条件分布律

$$P(Y=j|X=i)=\frac{p^2q^{j-2}}{pq^{i-1}}=pq^{j-i-1}(j=i+1,i+2\dots)$$

随机变量 Y 的边缘分布律为

$$P(Y=j)=C_{j-1}^{1}p^2q^{j-2}=(j-1)p^2q^{j-2}$$

所以在 Y=j (j=2，3，...) 的条件下，随机变量 X 的条件分布律

$$P(X=i|Y=j)=\frac{p^2q^{j-2}}{(j-1)p^2q^{j-2}}=\frac{1}{j-1}$$

连续型条件分布是怎么得来的？

关于连续型随机变量的条件分布函数的定义，很自然的想法就是把离散型随机变量分布函数的定义. 设 (X, Y) 是二维连续型随机变量，对于固定的 y 值 $P\left(X \leq x, Y = y_j\right) = 0$，$P\left(Y = y_j\right) = 0$, 如果直接套用离散型的公式，那么我们得到是 $\frac{0}{0}$，于是我们运用极限的定义。

$$\begin{aligned}
&\lim_{\Delta y \to 0+} \frac{P\left(X \leq x, y_j < Y \leq y_j + \Delta y\right)}{P\left(y_j < Y \leq y_j + \Delta y\right)} \\
=& \lim_{\Delta y \to 0+} \frac{F\left(x, y_j + \Delta y\right) - F\left(x, y_j\right)}{F_Y\left(y_j + \Delta y\right) - F\left(y_j\right)} \\
=& \lim_{\Delta y \to 0+} \frac{\left[F\left(x, y_j + \Delta y\right) - F\left(x, y_j\right)\right] / \Delta y}{\left[F_Y\left(y_j + \Delta y\right) - F_Y\left(y_j\right)\right] / \Delta y} \\
=& \frac{F_y'\left(x, y_j\right)}{f_Y\left(y_j\right)} = \frac{\int_{-\infty}^{x} f\left(u, y_j\right) du}{f_Y\left(y_j\right)} = \int_{-\infty}^{x} \frac{f\left(u, y_j\right)}{f_Y\left(y_j\right)} du
\end{aligned}$$

求导，我们就可以得到

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

例题 10

设数 X 在区间 (0,1) 均匀分布，当观察到 X=x (0<x<1) 时，数 Y 在区间 (x,1) 上随机地取值. 求 Y 的概率密度.

解答

依题意，X 具有概率密度

$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & others \end{cases}$$

对于任意给定的值 x (0<x<1)，在 X=x 的条件下，Y 的条件概率密度为

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{1-x}, & x < y < 1 \\ 0, & others \end{cases}$$

所以 X 和 Y 的联合密度为

$$f(x,y) = f_{Y|X}(y|x) f_X(x) = \begin{cases} \frac{1}{1-x}, & x < y < 1 \\ 0, & others \end{cases}$$

于是 Y 的概率密度为

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \begin{cases} -\ln(1-y), & 0 < y < 1 \\ 0, & others \end{cases}$$

3.3 随机变量的独立性

定义 3.8 设 X,Y 是两个随机变量，若对于任意的 $a,b(a<b);c,d(c<d)$, 事件 $\{a < X \leq b\}$ 和 $\{c < X \leq d\}$ 相互独立，即 $P\{a < X \leq b, c < Y \leq d\} = P\{a < X \leq b\}P\{c < Y \leq d\}$ 则称随机变量 X、Y 相互独立。

Tip X 和 Y 相互独立 $\Leftrightarrow F_X(x)F_Y(y)$

Tip (二维离散)X 和 Y 相互独立 $\Leftrightarrow P\{X = x_i, Y = y_j\} = P\{X = x_i\} P\{Y = y_j\}$

Tip (二维连续)X 和 Y 相互独立 $\Leftrightarrow p(x,y) = p_X(x)p_Y(y)$

例题 11

设 $(X,Y) \sim N(\mu_1, \sigma_X^1, \mu_2, \sigma_Y^2, \rho)$，若 X，Y 相互独立 $\Leftrightarrow \rho = 0$

解答

$\Leftarrow$ 由 $\rho=0$ 得:

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

很容易知道 $p(x,y)=p_X(x)p_Y(y)$, 所以 X，Y 独立。

$\Rightarrow$ 由于 $p(x,y)=p_X(x)p_Y(y)$ 那么

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}=\frac{1}{\sigma_1\sqrt{2\pi}}\frac{1}{\sigma_2\sqrt{2\pi}}\Rightarrow\sqrt{1-\rho^2}=1$$

推论 3.2 若连续型随机向量 $X_1,\ldots,X_n$ 的概率密度函数 $f(x_1,\ldots,x_n)$ 可表示为 n 个函数 $g_1,\ldots,g_n$ 之积，其中 g_i 只依赖于 x_i，即 $f(x_1,\ldots,x_n)=g_1(x_1)\ldots g_n(x_n)$ 则 $X_1,\ldots,X_n$ 相互独立, 且 X_i 的边缘密度 $f_i(x_i)$ 与 $g_i(x_i)$ 只相差一个常数因子.

推论 3.3 若 $X_1,\ldots,X_n$ 相互独立, 而 $Y_1=g_1(X_1,\ldots,X_m)$, $Y_2=g_2(X_{m+1},\ldots,X_n)$ 则 Y_1 与 Y_2 独立.

Tip 若两个随机变量相互独立, 且又有相同的分布, 不能说这两个随机变量相等

3.4 已知联合分布，求函数的分布

3.4.1 连续型随机变量和的分布

例题 12

设 X 和 Y 的联合密度为 p (x,y), 求 Z=X+Y 的密度

解答

$$F_Z(z)=\iint_{x+y\le z}p(x,y)dxdy$$

$$\begin{aligned}F_z(z)&=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{z-y}p(x,y)dx\right]dy\\&=\int_{-\infty}^{z-y}\left[\int_{-\infty}^{\infty}p(x,y)dy\right]dx\end{aligned}$$

$$p_Z(z)=F_Z'(z)=\int_{-\infty}^{\infty}p(z-y,y)dy$$

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

定理 3.1 (卷积公式) 当 X 和 Y 独立，设 (X,Y) 关于 X,Y 的边缘密度分别为 $p_X(x), p_Y(y)$,$Z = X + Y$ 那么

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

例题 13

若 X 和 Y 独立, 具有共同的概率密度 $p(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & others \end{cases}$ 求 Z=X+Y 的概率密度.

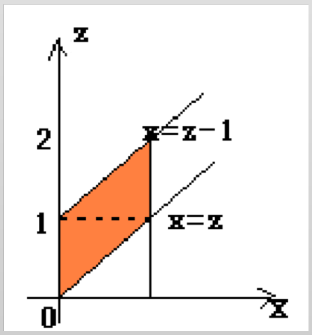
解答

确定积分区域

$$\begin{cases} 0 \le x \le 1 \\ 0 \le z - x \le 1 \end{cases}$$

由卷积公式可得:

$$p_Z(z) = \begin{cases} \int_0^z dx = z, & 0 \le z < 1 \\ \int_{z-1}^1 dx = 2 - z, & 1 \le z < 2 \\ 0, & others \end{cases}$$



例题 14

设随机变量 X_1 和 X_2 相互独立, 且均服从标准正态分布 $N \sim (0,1)$, 求 $Y = X_1 + X_2$ 的概率密度函数

解答

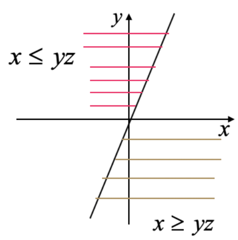
$$
\begin{aligned}
&p_Y(y)=\int_{-\infty}^{+\infty}p_1(y-x)p_2(x)dx=\tfrac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}}e^{-\frac{(y-x)^2}{2}}dx\\
&=\tfrac{1}{2\pi}e^{-\frac{y^2}{4}}\int_{-\infty}^{+\infty}e^{-\left(x-\frac{y}{2}\right)^2}dx\text{令}\tfrac{t}{\sqrt{2}}=x-\tfrac{y}{2}\left(\int_{-\infty}^{+\infty}e^{-u^2}du=\sqrt{\pi}\right)\\
&=\tfrac{1}{2\sqrt{2}\pi}e^{-\frac{y^2}{4}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=\tfrac{1}{2\sqrt{\pi}}e^{-\frac{y^2}{4}}
\end{aligned}
$$

推论 3.4 两个独立的正态分布的随机变量的和仍服从正态分布. 即:$X_1\sim N(\mu_1,\sigma_1^2),X_2\sim N(\mu_2,\sigma_2^2)$,$X_1,X_2$ 独立，则 $X_1+X_2\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$

推论 3.5 若 $X_i\sim N(\mu_i,\sigma_i^2),(i=1,2,\dots n),X_1,X_2,\dots X_n$ 相互独立, 实数 $a_1,a_2,\dots,a_n$ 不全为零, 则

$$\sum_{i=1}^{n}a_iX_i\sim N\left(\sum_{i=1}^{n}a_i\mu_i,\sum_{i=1}^{n}a_i^2\sigma_i^2\right)$$

3.4.2 连续型随机变量商的分布



计算随机变量 $Z=\frac{X}{Y}$ 的密度函数。

$$
\begin{aligned}
&F_Z(z)=P\{Z\le z\}=P\left\{\tfrac{X}{Y}\le z\right\}\\
&=\iint\limits_{\frac{x}{y}\le z}p(x,\quad y)dxdy\\
&=\iint\limits_{\frac{x}{y}\le z,y>0}p(x,\quad y)dxdy+\iint\limits_{\frac{x}{y}\le z,y<0}p(x,\quad y)dxdy\\
&=\iint\limits_{x\le zy,y>0}p(x,\quad y)dxdy+\iint\limits_{x\ge yz,y<0}p(x,\quad y)dxdy\\
&=\int_0^{+\infty}dy\int_{-\infty}^{zy}p(x,\quad y)dx+\int_{-\infty}^{0}dy\int_{zy}^{+\infty}p(x,\quad y)dx
\end{aligned}
$$

对于第一个积分，我们做变量代换 $x=uy$, 那么当 $x=zy$ 时，$u=z$，当 $x\to-\infty$, 由于 $y>0$ 所以 $u\to-\infty$;

$$
\begin{aligned}
&\int_0^{+\infty}dy\int_{-\infty}^{xy}p(x,y)dx=\int_0^{+\infty}dy\int_{-\infty}^{z}p(uy,\quad y)ydu\\
&=\int_{-\infty}^{z}du\int_0^{+\infty}yp(uy,\quad y)dy=\int_{-\infty}^{z}du\int_0^{+\infty}|y|p(uy,\quad y)dy
\end{aligned}
$$

对于第二个积分做同样的标量代换，我们可以得到：

$$p_Z(z)=\int_{-\infty}^{+\infty}|y|p(zy,\quad y)dy$$

推论 3.6 设 (X,Y) 是二维连续型随机变量，其联合密度函数为 $p(x,y)$，令 $Z=X-Y$，则

$$p_Z(z)=\int_{-\infty}^{+\infty}p(z+y,\quad y)dy$$

推论 3.7 设 (X,Y) 是二维连续型随机变量，其联合密度函数为 $p(x,y)$，令 $Z=XY$，则

$$p_Z(z)=\int_{-\infty}^{+\infty}p\left(x,\quad \frac{z}{x}\right)\frac{1}{|x|}dx=\int_{-\infty}^{+\infty}p\left(\frac{z}{y},y\right)\frac{1}{|y|}dy$$

3.4.3 随机变量的 max 和 min 分布

M=max(X,Y) 不大于 z 等价于 X 和 Y 都不大于 z 故有 $P(M\le z)=P(X\le z,Y\le z)$ 又由于 X 和 Y 独立

$$F_M(z)=P(M\le z)=P(X\le z,Y\le z)=P(X\le z)P(Y\le z)\Rightarrow F_M(z)=F_X(z)F_Y(z)$$

类似的，可得 N=min(X,Y) 的分布函数为

$$\begin{aligned}\boldsymbol{F}_N(z)&=\boldsymbol{P}(\boldsymbol{N}\le z)=\mathbf{1}-\boldsymbol{P}(\boldsymbol{N}>z)\\&=\mathbf{1}-\boldsymbol{P}(\boldsymbol{X}>z,\boldsymbol{Y}>z)\\&=\mathbf{1}-\boldsymbol{P}(\boldsymbol{X}>z)P(\boldsymbol{Y}>z)\end{aligned}$$

即有 $F_N(z)=1-[1-F_X(z)][1-F_Y(z)]$

更近一步：M=max$(X_1,\ldots,X_n)$ 的分布函数为:

$$F_M(z)=F_{X_1}(z)\ldots F_{X_n}(z)$$

N=min$(X_1,\ldots,X_n)$ 的分布函数为:

$$F_N(z)=1-[1-F_{X_1}(z)]\ldots[1-F_{X_n}(z)]$$

3.5 二维随机变量的推广

- n 维随机变量 $(X_1, X_2, \ldots, X_n)$ 的分布函数 $F(x_1, x_2, \ldots, x_n) = P\{X_1 \leq x_1, \ldots, X_n \leq x_n\}$
- 概率密度函数满足 $F(x_1, x_2 \ldots, x_n) = \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n-1}} \cdots \int_{-\infty}^{x_1} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n$
- 边缘分布函数。$F_{X_1}(x_1) = F(x_1, \infty, \ldots, \infty)$ 称为 n 维随机变量 $(X_1, X_2 \ldots, X_n)$ 关于 X_1 的边缘分布函数，$F_{X_1 X_2}(x_1, x_2) = F(x_1, x_2, \ldots, \infty)$ 称为 n 维随机变量 $(X_1, X_2 \ldots, X_n)$ 关于 (X_1, X_2) 的边缘分布函数
- 边缘密度函数。
- 相互独立性。若对所有的 $x_1, \ldots, x_n$ 有 $F(x_1, x_2, \cdots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n)$ 则称 $X_1, \ldots, X_n$ 是相互独立的。

4 随机变量的数字特征

4.1 一维随机变量的数字特征

定义 4.1 (离散型随机变量的数学期望) $P\{X=x_k\}=p_k$，若级数 $\sum\limits_{k=1}^{\infty}$ 绝对收敛，则称级数 $\sum\limits_{k=1}^{\infty}$ 为随机变量 X 的数学期望。

重要分布的数学期望

- 0-1 分布

$$E(X)=1\times p+0\times(1-p)=p$$

- 二项分布 B(n,p)

$$\begin{aligned}
E(X)&=\sum_{k=1}^{n}k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}\\
&=\sum_{k=1}^{n}\frac{n!}{(k-1)!(n-k)!}p^k(1-p)^{n-k}\\
&=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}(1-p)^{n-1-(k-1)}\\
&\overset{l=k-1}{=}np\sum_{l=0}^{n-1}C_{n-1}^{l}p^l(1-p)^{n-1-l}=np
\end{aligned}$$

- 泊松分布

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

定义 4.2 (连续型随机变量的数学期望) 设 X 是连续型随机变量，其密度函数为 $f(x)$, 如果积分 $\int_{-\infty}^{\infty}xf(x)dx$ 绝对收敛, 则称此积分值为 X 的数学期望, 即

$$E(x)=\int_{-\infty}^{\infty}xf(x)dx$$

连续型随机变量的数学期望公式是怎么得来的？

设 X 是连续型随机变量, 其密度函数为 f (x), 在数轴上取很密的分点 $x_0<x_1<x_2<\ldots$, 由于 x_i 与 x_{i+1} 很接近, 所以区间 $[x_i,x_{i+1})$ 中的值可以用 x_i 来近似代替.$\lim_{\lambda\to 0}\sum_k x_kf(x_k)\Delta x_k=\int_{-\infty}^{+\infty}xf(x)dx$

重要分布的数学期望 (续)

- 均匀分布 U(a,b)

$$E(X)=\int_a^b \frac{x}{b-a}dx=\frac{a+b}{2}$$

- 指数分布

$$\begin{aligned}E(X)&=\int_0^\infty x\lambda e^{-\lambda x}dx=-\int_0^\infty x de^{-\lambda x}\\&=-\left.xe^{-\lambda x}\right|_0^\infty+\int_0^\infty e^{-\lambda x}dx=\frac{1}{\lambda}\end{aligned}$$

- 正态分布 $N(\mu,\sigma^2)$

$$\begin{aligned}&E(X)=\textstyle\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx\\&\overset{t=\frac{x-\mu}{\sigma}}{=}\textstyle\int_{-\infty}^{\infty}\frac{\sigma t+\mu}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\sigma dt;\\&=\mu\end{aligned}$$

数学期望的性质

- E(c)=c,c 为常数;
- E(cX)=cE(X)
- E(X+Y)=E(X)+E(Y);

$$\begin{aligned}&E(X+Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dxdy\\&=\int_{-\infty}^{\infty}\int_x xf(x,y)dxdy+\int_{-\infty}^{\infty}yf(x,y)dxdy\\&=\int_{-\infty}^{\infty}x\left[\int_{-\infty}^{\infty}f(x,y)dy\right]dx+\int_{-\infty}^{\infty}y\left[\int_{-\infty}^{\infty}f(x,y)dx\right]dy\\&=\int_{-\infty}^{\infty}xf_X(x)dx+\int_{-\infty}^{\infty}yf_Y(y)dy=E(X)+E(Y)\end{aligned}$$

- 若 X 与 Y 独立，则 E(XY)=E(X)E(Y)

$$\begin{aligned}E(XY)&=\int_{-\infty}^{\infty}\int_{x=\infty}^{\infty}xyf(x,y)dxdy\\&=\int_{-\infty}^{\infty}\int_{x=\infty}^{\infty}xyf_X(x)f_Y(y)dxdy\\&=\int_{-\infty}^{\infty}xf_X(x)dx\int_{-\infty}^{\infty}yf_Y(y)dy=E(X)E(Y)\end{aligned}$$

随机变量的函数的数学期望

设 Y 是随机变量 X 的函数:Y=g (X)

- 当 X 为离散型时, 它的分布率为 $P(X=x_k)=p_k(k=1,2\ldots)$, 若 $\sum\limits_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛, 则有

$$\boldsymbol{E(Y)=E[g(X)]=\sum_{k=1}^{\infty} g\left(x_k\right) p_k}$$

- 当 X 为连续型时, 它的密度函数为 f(x). 若 $\int\limits_{-\infty}^{\infty} g(x)f(x)dx$ 绝对收敛，则有

$$\boldsymbol{E(Y)=E[}g(\boldsymbol{X})\boldsymbol{]}=\int_{-\infty}^{+\infty} g(x)\boldsymbol{f(x)dx}$$

4.2 二维随机变量的数字特征

定理 4.1 设二维离散型随机变量 *(X,Y)* 的联合概率分布为 $P\left\{X=x_i,Y=y_j\right\}=p_{ij},i,j=1,2\ldots$ 则

$$E(X)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} x_i p_{ij},\quad E(Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} y_j p_{ij}$$

定理 4.2 设二维连续型随机变量 *(X,Y)* 的概率密度函数为 $f(x,y)$, 则有

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\mathrm{d}x\mathrm{d}y, E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x,y)\mathrm{d}x\mathrm{d}y$$

证明

$$\begin{aligned}
E(X)&=\int_{-\infty}^{+\infty} xf_X(x)\mathrm{d}x=\int_{-\infty}^{+\infty} x\left[\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y\right]\mathrm{d}x\\
&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\mathrm{d}x\mathrm{d}y\\
E(Y)&=\int_{-\infty}^{+\infty} yf_y(y)\mathrm{d}y=\int_{-\infty}^{+\infty} y\left[\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x\right]\mathrm{d}y\\
&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x,y)\mathrm{d}x\mathrm{d}y
\end{aligned}$$

定义 4.3 如果 *(X,Y)* 为离散型随机向量, 其联合概率分布为 $P\{X=x_iY=y_j\}=p_{ij}i,j=1,2,3,\ldots,$, 如果 $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|g(x_i,y_j)|p_{ij}<+\infty$ 设二维随机向量（X，Y）为连续型随机变量，它的联合概率密

度为 $f(x,y)$, 若 $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|g(x,y)|f(x,y)dxdy$ 收敛, 则 $Z=g(X,Y)$ 的数学期望为

$$E(Z)=E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy$$

例题 15

为普查某种疾病,n 个人需验血. 验血方案有如下两种:

- (1) 将每个人的血分别化验，共需化验 n 次;
- (2) 分组化验，按 k 个人一组进行分组，即把 k 个人的血混在一起化验, 若结果为阴性，说明这 k 个人的血都呈阴性反应, 则只需化验一次；若为阳性，则对 k 个人的血逐个化验，找出患病者，此时 k 个人的血需化验 k+1 次.

你认为应该选择哪一方案，并说明你的理由. 如果选取第二种方案，应如何进行分组?

解答

下面我们计算第二种方案所需化验的次数. 根据已知条件，各人的血呈阴性反应的概率为 q=1-p. 因而 k 个人的血呈阴性反应的概率为 q^k , 呈阳性反应的概率为 $1-q^k$.
设第 i 组化验次数 X_i，则 X_i 的分布律为 $P(x_i=1)=q^k, P(x_i=k+1)=1-q^k$

$$E(X_i)=1\times q^k+(k+1)\times(1-q^k)$$

那么 n 个人的平均化验次数为

$$E(X)=E(X_1+X_2+\cdots+X_{\frac{n}{k}})=n(1-q^k+\frac{1}{k})$$

然后比较就可以了。

4.3 方差

定义 4.4 设 X 是一个随机变量, 若 $E[(X-E(X)]^2$ 存在, 称 $E[(X-E(X)]^2$ 为 X 的方差, 记为 $D(X)$

$$D(X)=\begin{cases}\sum_{k=1}^{\infty}[x_k-E(X)]^2p_k\\ \int_{-\infty}^{\infty}[x-E(X)]^2f(x)dx\end{cases}$$

定理 4.3

$$D(X)=E(X^2)-[E(X)]^2$$

证明

$$
\begin{aligned}
D(X) &= E[X - E(X)]^2 \\
&= E\left\{X^2 - 2XE(X) + [E(X)]^2\right\} \\
&= E\left(X^2\right) - 2[E(X)]^2 + [E(X)]^2 \\
&= E\left(X^2\right) - [E(X)]^2
\end{aligned}
$$

重要分布的方差

- 0-1 分布

$$E(X)=\mathbf{0}\times(1-p)+1\times p=p$$

$$E\left(X^2\right)=0^2\times(1-p)+1^2\times p=p$$

$$D(X)=E\left(X^2\right)-[E(X)]^2=p-p^2=p(1-p)$$

- 二项分布 B(n,p)

$$\begin{aligned}E\left(X^2\right)&=E[X(X-1)+X)]=E[X(X-1)]+E(X)\\&=\sum_{k=0}^{n}k(k-1)C_n^k p^k q^{n-k}+np\\&=\sum_{k=0}^{n}k(k-1)\frac{n!}{k!(n-k)!}p^k q^{n-k}+np\\&=n(n-1)p^2\sum_{k=2}^{n}\frac{(n-2)!}{(k-2)!(n-k)!}p^{k-2}q^{n-k}+np\\&=n(n-1)p^2(p+q)^{n-2}+np=n(n-1)p^2+np\end{aligned}$$

$$D(X)=E\left(X^2\right)-[E(X)]^2=n(n-1)p^2+np-n^2p^2=np(1-p)$$

- 泊松分布

$$\begin{aligned}\boldsymbol{E}\left(\boldsymbol{X}^2\right)&=\boldsymbol{E}[\boldsymbol{X}(\boldsymbol{X}-\mathbf{1})+\boldsymbol{X}]=\boldsymbol{E}[\boldsymbol{X}(\boldsymbol{X}-\mathbf{1})]+\boldsymbol{E}(\boldsymbol{X})\\&=\sum_{k=0}^{\infty}k(\boldsymbol{k}-1)\frac{\lambda^k\boldsymbol{e}^{-\lambda}}{k!}+\lambda=\lambda^2\boldsymbol{e}^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(\boldsymbol{k}-2)!}+\lambda\\&=\lambda^2\boldsymbol{e}^{-\lambda}\boldsymbol{e}^{\lambda}+\lambda=\lambda^2+\lambda\end{aligned}$$

$$\Rightarrow D(X)=E(X^2)-[E(X)]^2=\lambda$$

- 均匀分布 U(a,b)

$$\begin{aligned}D(X)&=E\left(X^2\right)-E(X)^2\\&=\int_a^b x^2\frac{1}{b-a}dx-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}\end{aligned}$$

重要分布的方差 (续)

- 指数分布 $E(\lambda)$

$$p(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\leq 0\\ 0 & x\leq 0\end{cases}$$
$$E(X)=\int_{-\infty}^{+\infty} xp(x)dx=1/\lambda$$
$$E\left(X^2\right)=\int_{-\infty}^{+\infty} x^2p(x)dx=2/\lambda^2$$
$$\Rightarrow D(X)=1/\lambda^2$$

- 正态分布 $X\sim N(\mu,\sigma^2)$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},(-\infty<x<+\infty)$$
$$E(X)=\mu$$
$$\begin{aligned}D(X)=E[X-E(X)]^2&=\int_{-\infty}^{+\infty}(x-\mu)^2\cdot\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(-\alpha-\mu)^2}{2\sigma^2}}\mathrm{d}x\\ &=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t\quad\left(t=\frac{x-\mu}{\sigma}\right)\\ &=\frac{\sigma^2}{\sqrt{2\pi}}\left[t\left(-\mathrm{e}^{-\frac{t^2}{2}}\right)\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t\right]\\ &=\frac{\sigma^2}{\sqrt{2\pi}}\cdot\sqrt{2\pi}=\sigma^2\end{aligned}$$

方差的性质

- 设 C 是常数, 则有 D(C)=0
- 设 X 是一个随机变量, C 是常数, 则有 $D(CX)=C^2D(X), D(X+C)=D(X)$

$$\begin{aligned}\boldsymbol{D}(\boldsymbol{CX}) &= \boldsymbol{E}\left\{[\boldsymbol{CX}-\boldsymbol{E}(\boldsymbol{CX})]^2\right\} \\ &= \boldsymbol{C}^2\boldsymbol{E}\left\{[\boldsymbol{X}-\boldsymbol{E}(\boldsymbol{X})]^2\right\} = \boldsymbol{C}^2\boldsymbol{D}(\boldsymbol{X}) \\ \boldsymbol{D}(\boldsymbol{X}+\boldsymbol{C}) &= \boldsymbol{E}\left\{[(\boldsymbol{X}+\boldsymbol{C})-\boldsymbol{E}(\boldsymbol{X}+\boldsymbol{C})]^2\right\} \\ &= \boldsymbol{E}\left\{[\boldsymbol{X}-\boldsymbol{E}(\boldsymbol{X})]^2\right\} \\ &= \boldsymbol{D}(\boldsymbol{X})\end{aligned}$$

- 设 X,Y 是两个随机变量, 则有 $D(X+Y) = D(X)+D(Y)+2E\left\{(X-E(X))(Y-E(Y))\right\}$

$$\begin{aligned}&D(X+Y)=E\left\{[(X+Y)-E(X+Y)]^2\right\} \\ &=E\left\{[(X-E(X))+(Y-E(Y))]^2\right\} \\ &=E[X-E(X)]^2+E[Y-E(Y)]^2+2E\{[X-E(X)][Y-E(Y)]\} \\ &=D(X)+D(Y)+2E\{[X-E(X)][Y-E(Y)]\}\end{aligned}$$

- $D(X)=0$ 的充要条件是 X 以概率 1 取常数 C。

推论 4.1 若 X, Y 独立，则 $D(X+Y)=D(X)+D(Y)$

推论 4.2 若 $X_1,\ldots,X_n$ 独立，则

$$D(\sum_{i=1}^n X_i)=\sum_{i=1}^n D(X_i)$$

定理 4.4 (切比雪夫不等式) 设随机变量 X，其数学期望 $E(X)$, 方差存在，假定 $D(X)=\sigma^2$，则对于任意正数 ε, 有不等式

$$\boldsymbol{P}\{|\boldsymbol{X}-\boldsymbol{E}(\boldsymbol{X})\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}$$

证明

$$\begin{aligned}\mathbf{P}\{|\mathbf{X}-\mu|\geq\varepsilon\} = \int_{|x-\mu|\geq\varepsilon} f(x)dx \quad &\leq \int_{|x-\mu|\geq\varepsilon}\frac{(x-\mu)^2}{\varepsilon^2}f(x)dx \\ \leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2f(x)dx &= \frac{\mathbf{D}(\mathbf{X})}{\varepsilon^2}=\frac{\sigma^2}{\varepsilon^2}\end{aligned}$$

例题 16

在每次试验中，事件 A 发生的概率为 0.75, 利用切比雪夫不等式求：n 需要多么大时，才能使得在 n 次独立重复试验中, 事件 A 出现的频率在 0.74 ∼ 0.76 之间的概率至少为 0.90?

解答

设 X 为 n 次试验中，事件 A 出现的次数，则 $X \sim B(n, 0.75)$,

$$E(X) = 0.75n, D(X) = 0.75 \times 0.25n = 0.1875n$$

所求为满足 $P(0.74 < \frac{X}{n} < 0.76) \geq 0.90$ 的最小的 n。

$$\begin{aligned} &P(0.74n < X < 0.76n) \\ =&P(-0.01n < X - 0.75n < 0.01n) \\ =& P\left\{|X - E(X)|\right\} < 0.01n \end{aligned}$$

在切比雪夫不等式中取 $\varepsilon = 0.01n$，则

$$\begin{aligned} &P\left(0.74 < \tfrac{X}{n} < 0.76\right) = P\{|X - E(X)| < 0.01n\} \\ &\geq 1 - \tfrac{D(X)}{(0.01n)^2} = 1 - \tfrac{0.1875n}{0.0001n^2} = 1 - \tfrac{1875}{n} \end{aligned}$$

取 $1 - \frac{1875}{n} \geq 0.9$ 得 $n \geq 18750$

4.4 协方差与相关系数

定义 4.5 量 $E\{[X - E(X)][Y - E(Y)]\}$ 称为随机变量 X 和 Y 的协方差, 记为 *Cov(X,Y)* ，即

$$\begin{aligned} &\operatorname{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\} \\ =&E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ =&E(XY) - E(X)E(Y) \end{aligned}$$

协方差的性质

- Cov(X,Y)= Cov(Y,X)
- Cov(aX,bY) = ab Cov(X,Y) a,b 是常数
- Cov($X_1 + X_2$,Y)= Cov(X_1,Y) + Cov(X_2,Y)
- 若 X 与 Y 独立，Cov(X,Y)= 0
- $Cov(X, X) = E(X^2) - E(X)^2 = D(X)$
- D(X+Y)= D(X)+D(Y)+ 2Cov(X,Y)

为什么要引进相关系数

协方差的数值在一定程度上反映了 X 与 Y 相互间的联系, 但它受 X 与 Y 本身数值大小的影响. 如令 X*=kX,Y*=kY, 这时 X* 与 Y* 间的相互联系和 X 与 Y 的相互联系应该是一样的, 但是

$$Cov(kX, kY) = k^2Cov(X, Y)$$

为了克服这一缺点，对协方差进行标准化，这就引入了相关系数.

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}} \quad Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}}$$

定义 4.6 设 *D(X)>0, D(Y)>0,* 记随机变量 *X* 和 *Y* 的相关系数为

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)D(Y)}}$$

相关系数的性质

- 随机变量 X 和 Y 的相关系数满足 $|\rho_{XY}| \leq 1$.

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}} \quad Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}}$$

- $|\rho_{XY}| = 1$ 的充要条件是, 存在常数 a,b 使得 P{Y=aX+b}=1
- 若 X 与 Y 相互独立, 则 ρ_{XY}=0.

Tip 相互独立一定不相关，但是不相关不一定独立。

不相关但是不独立的例子

设 X 服从 (-1/2, 1/2) 内的均匀分布, 而 Y=cos X,Cov(X,Y)=0

Tip 需要指出的是，相关系数有一个明显的缺点，即它接近于 1 的程度与数据组数 n 相关，这容易给人一种假象。因为，当 n 较小时，相关系数的波动较大，对有些样本相关系数的绝对值易接近于 1；当 n 较大时，相关系数的绝对值容易偏小。特别是当 n=2 时，相关系数的绝对值总为 1。因此在样本容量 n 较小时，我们仅凭相关系数较大就判定变量 x 与 y 之间有密切的线性关系是不妥当的。

4.5 随机变量的其它数字特征

设 X 和 Y 是随机变量

- 若 $E(X^k)$(k=1,2,$\cdots$) 存在, 则称它为 X 的 k 阶原点矩.
- 若 $E[X-E(X)]^k$ (k=1,2,$\cdots$) 存在, 则称它为 X 的 k 阶中心矩.
- 若 $E(X^kY^l)$ (k,l=1,2,$\cdots$) 存在, 则称它为 X 和 Y 的 k+l 阶混合矩.
- 若 $E[X-E(X)]^k[Y-E(Y)]^l$ (k,l=1,2,$\cdots$) 存在, 则称它为 X 和 Y 的 k+l 阶混合中心矩.

定义 4.7 若随机变量 X 的方差存在，则称 $\frac{\sqrt{D(X)}}{E(X)}$ 为 X 的变异系数，记为 $CV(X)=\frac{\sqrt{D(X)}}{E(X)}$

定义 4.8 设连续型随机变量 X 的密度函数为 *f(x)*，*0<x<1*，则称满足 $\int_{-\infty}^{x_p} f(x)dx=p$ 的实数 x_p 为 X 的下侧 p 分位数，称满足 $\int_{y_p}^{\infty} f(x)dx=p$ 的实数 y_p 为 X 的上侧 p 分位数。

例题 17

把数字 1,2,$\cdots$,n 任意地排成一列，如果数字 k 恰好出现在第 k 个位置上，则称为一个巧合，求巧合个数的数学期望.

解答

设巧合个数为 X, 引入

$$X_k \begin{cases} 1 & \text{数字 k 恰好出现在第 k 个位置上} \\ 0 & \text{否则} \end{cases}$$

则 $X = \sum_{k=1}^{n} X_k$, 由于 $E(X_k) = \frac{(n-1)!}{n!} = \frac{1}{n}$, 故

$$E(X) = \sum_{k=1}^{n} E(X_k) = n \times \frac{1}{n} = 1$$

5 大数定理及中心极限定理

5.1 大数定律

定义 5.1 设 $\{X_n\}$ 为一随机变量序列，X 为一随机变量或常数，若对任意 $\epsilon > 0$，有

$$\lim_{n\to\infty} P\{|X_n - X| < \varepsilon\} = 1$$

则称 $\{X_n\}$ 依概率收敛于 X

定理 5.1 (伯努利大数定律) 设 n_A 是 n 次独立重复试验中事件 A 发生的次数, p 是每次试验中 A 发生的概率，则 $\forall\epsilon > 0$ 有

$$\lim_{n\to\infty} P\left(\frac{n_A}{n} - p \geq \varepsilon\right) = 0$$
$$\lim_{n\to\infty} P\left(\frac{n_A}{n} - p < \varepsilon\right) = 1$$

> **Tip** 当 n 很大时, 事件发生的频率与概率有较大偏差的可能性很小. 在实际应用中, 当试验次数很大时, 便可以用事件发生的频率来代替事件的概率.

定义 5.2 (切比雪夫大数定律) 设随机变量序列 $X_1, X_2, \ldots, X_n$ 相互独立，且具有相同的数学期望和方差

$$E(X_k) = \mu, D(X_k) = \sigma^2$$

则 $\forall\epsilon > 0$ 有

$$\lim_{n\to\infty} P\left(|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu| \geq \varepsilon\right) = 0$$

证明

通过切比雪夫不等式来证明。

$$P\{|Y_n - E(Y_n)| < \varepsilon\} \geq 1 - \frac{D(Y_n)}{\varepsilon^2}$$
$$E(Y_n) = \frac{1}{n}\sum_{k=1}^{n} E(X_k) = \mu$$
$$D(Y_n) = \frac{1}{n^2}\sum_{k=1}^{n} D(X_k) = \frac{\sigma^2}{n}$$
$$P\{|Y_n - \mu| < \varepsilon\} \geq 1 - \frac{\sigma^2}{n\varepsilon^2}$$
$$\lim_{n\to\infty} P\{|Y_n - \mu| < \varepsilon\} = 1$$

注

- $X_1, X_2, \ldots, X_n, \ldots$ 不一定有相同的数学期望与方差，可设 $E(X_k) = \mu_k, D(X_k) = \sigma_k^2 \leq \sigma^2 (k = 1, 2 \ldots)$
- $X_1, X_2 \ldots, X_n, \ldots$ 相互独立的条件可以去掉，代之以

$$\frac{1}{n^2} D\left(\sum_{k=1}^{n} X_k\right) \stackrel{n\to\infty}{\to} 0$$

定理 5.2 (辛钦大数定律) 设随机变量序列 $X_1, X_2, \ldots$ 相互独立，服从同一分布，具有数学期 $E(X_i) = \mu, i = 1, 2, \ldots$ 则对于任意正数 ϵ, 有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1$$

Tip 辛钦大数定律比切比雪夫大数定律的条件更弱，不要求随机变量的方差存在.

5.2 中心极限定理

定理 5.3 (中心极限定理) 如果一个随机变量是由大量相互独立的随机因素的综合影响所造成，而每一个别因素对这种综合影响中所起的作用不大. 则这种随机变量一般都服从或近似服从正态分布. 在概率论中，习惯于把和的分布收敛于正态分布这一类定理都叫做中心极限定理.

定理 5.4 (林德贝格-列维中心极限定理) 设随机变量 $X_1, X_2, \ldots$ 相互独立，服从同一分布，且具有数学期望和方差 $E(X_k) = \mu, D(X_k) = \sigma^2 > 0 (k = 1, 2 \ldots)$ 且随机变量之和 $\sum_{k=1}^{n} X_k$ 的标准化变量 $Y_n = \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$，对于任意 x，满足

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

注

- 独立同分布的随机变量之和 $\sum_{k=1}^{n} X_k$，当 n 充分大的时候，随机变量之和与其标准化变量分别有 $\sum_{k=1}^{n} \sim N(n\mu, n\sigma^2)$
- 独立同分布中心极限定理的另一种形式可写成 $\bar{X} \sim N(\mu, \sigma^2/n), \bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$
- 这个等价不是严格意义上的等价，而是近似。

定理 5.5 (李雅普诺夫 (Liapounov) 定理) 设随机变量 $X_1, X_2, \ldots$ 相互独立，服从同一分布，且具有数学期望和方差 $E(X_k) = \mu_k, D(X_k) = \sigma_k^2 > 0 (k = 1, 2 \ldots)$, 记 $B_n^2 = \sum_{k=1}^{n} \sigma_k^2$，若存在正数 δ，使得当 $n \to \infty$ 时，

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} E\left\{X_k - \mu_k^{2+\delta}\right\} \to 0$$

则随机变量之和 $\sum_{k=1}^{n} X_k$ 的标准化变量 $\frac{\sum_{k=1}^{n} X_k - E(\sum_{k=1}^{n} X_k)}{\sqrt{D(\sum_{k=1}^{n} X_k)}} = \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_n}$，对于任意 x，满足

$$\lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_n} \le x\right\} = \int_{\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\mathrm{t}^2/2} \mathrm{dt} = \Phi(x)$$

定理 5.6 (棣莫弗——拉普拉斯中心极限定理) 设随机变量 $\eta_n (n = 1, 2, \ldots)$ 服从参数为 n,p 的二项分布，则对任意 x，有

$$\begin{aligned} &\lim_{n\to\infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \le x\right\} \\ &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x) \end{aligned}$$

例题 18

100 个独立工作 (工作的概率为 0.9) 的部件组成一个系统，求系统中至少有 85 个部件工作的概率.

解答

记 $Y = X_1 + X_2 + \cdots + X_100$, 那么 $E(Y) = 90, Var(Y) = 9$

$$P\{Y \ge 85\} \approx 1 - \Phi\left(\frac{85 - 0.5 - 90}{\sqrt{9}}\right) = 0.966$$

例题 19

有 200 台独立工作 (工作的概率为 0.7) 的机床，每台机床工作时需 15kw 电力. 问共需多少电力, 才可有 95% 的可能性保证正常生产?

解答

$X_i = 1$ 表示第 i 台机床正常工作, 反之记为 $X_i = 0$. 记 $Y = X_1 + X_2 + \cdots + X_200$, 那么 $E(Y) = 140, Var(Y) = 42$，设供电量为 y，则

$$P\{15Y \le y\} \approx \Phi\left(\frac{y/15 + 0.5 - 140}{\sqrt{42}}\right) \ge 0.95$$

例题 20

用调查对象中的收看比例 k/n 作为某电视节目的收视率 p 的估计。要有 90%的把握，使 k/n 与 p 的差异不大于 0.05，问至少要调查多少对象?

解答

Y_n 表示 n 个调查对象中收看此节目的人数，则 Y_n 服从 b(n, p) 分布，k 为 Y_n 的实际取值。

$$P\left(|Y_n/n - p| < 0.05\right) \approx 2\Phi(0.05\sqrt{n/p(1-p)}) - 1 \ge 0.90$$

大数定律和中心极限定律的区别

设 $\{X_n\}$ 为独立同分布随机变量序列，且

$$E(X_i) = \mu, D(X_i) = \sigma^2 > 0$$

由大数定律可以知道，

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1$$

大数定律没有给出 $\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\}$ 的表达式，但是保证了它的极限是 1，但是中心极限定律是给出了表达式了的。

6 样本及抽样分布

定义 6.1 研究对象的全体称为总体，总体中每个对象称为个体. 为推断总体分布及各种特征, 按一定规则从总体中抽取若干个体进行观察试验以获得有关总体的信息. 所抽取的部分个体称为样本. 样本中所包含的个体数目称为样本容量.

定义 6.2 简单随机样本: 满足独立同分布（*i.i.d*）且和所考察的总体具有相同的分布。

定义 6.3 $(x_1,\ldots,x_n)$ 称为样本值，若总体 X 的分布函数为 $F(x)$，则其样本的联合分布函数为 $F(x_1,x_2,\ldots,x_n)=F(x_1)F(x_2)\ldots F(x_n)=\Pi_{i=1}^{n}F(x_i)$, 若总体 X 的概率密度函数为 $p(x)$，则其样本的联合密度函数为 $p(x_1,x_2,\ldots,x_n)=p(x_1)p(x_2)\ldots p(x_n)=\Pi_{i=1}^{n}p(x_i)$,

定义 6.4 (样本分布函数) 设总体的简单随机样本 $X=(X_1,X_2,\ldots,X_n)$ 的观察值为 $x=(x_1,x_2,\ldots,x_n)$ 从小到大排序为 $x_{(1)}\le x_{(2)}\le\ldots x_{(n)}$

$$F_n(x)=\begin{cases}0, x<x_{(1)}\\1/n, x_{(1)}\le x<x_{(2)}\\\cdots\cdots\cdots\cdots\\k/n, x_{(k)}\le x<x_{(k+1)}\\\cdots\cdots\cdots\cdots\cdots\\1, x\ge x_{(n)}\end{cases}$$

称 $F_n(x)$ 为样本分布函数或总体的经验分布函数。

定理 6.1 (格列纹科定理) 设 $x_1,x_2,\ldots,x_n$ 是取自体分布函数为 *F(x)* 的样本, $F_n(x)$ 是其经验分布函数，当 $n\to\infty$ 时，有

$$\lim_{n\to\infty}P\left\{\sup_{x\in R}|F_n(x)-F(x)|\le\varepsilon\right\}=1,\forall\varepsilon>0$$
$$P\left\{\lim_{n\to\infty}\sup_{x\in R}|F_n(x)-F(x)|=0\right\}=1$$

定义 6.5 (箱线图) 用来反映一组或多组连续型定量数据分布的中心位置和散布范围。在箱线图中，箱子的中间有一条线，代表了数据的中位数。箱子的上下底，分别是数据的上四分位数（Q_3）和下四分位数（Q_1），这意味着箱体包含了 *50%* 的数据。因此，箱子的高度在一定程度上反映了数据的波动程度。上下边缘则代表了该组数据的最大值和最小值。有时候箱子外部会有一些点，可以理解为数据中的“异常值”。

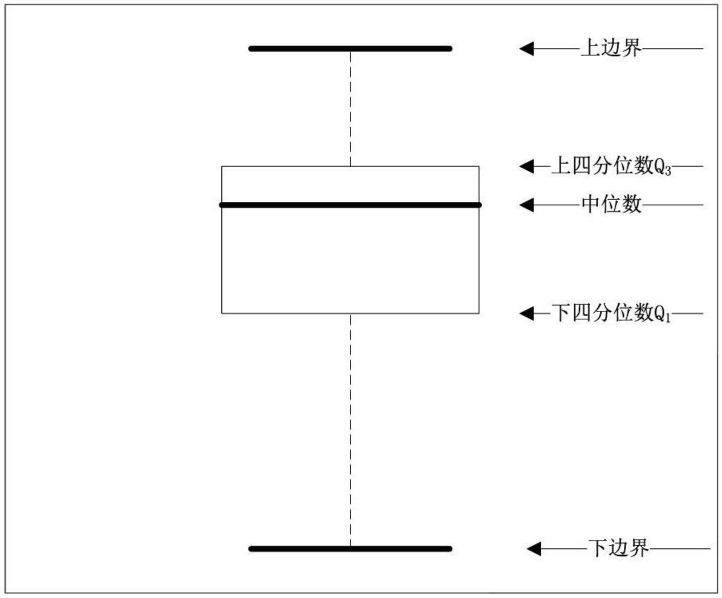


图 1　箱线图示例

定义 6.6 (统计量) 设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的一个样本，$g(X_1, \ldots, X_n)$ 是 $X_1, X_2, \ldots, X_n$ 的函数，若 g 不含未知参数，则称 $g(X_1, \ldots, X_n)$ 是一个统计量。

常见统计量

- 样本均值

$$\overline{\boldsymbol{X}} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- 样本方差

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

- 样本 k 阶原点矩

$$\boldsymbol{A_k} = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

- 样本 k 阶中心矩

$$\boldsymbol{B_k} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{\boldsymbol{X}})^k$$

性质

如果总体 X 的期望为 μ，方差为 σ^2，则

$$(1)\ E(\bar{X})=E(X)=\mu \quad (2) D(\bar{X})=\frac{D(X)}{n}=\frac{\sigma^2}{n}$$
$$(3)\ E\left(S^2\right)=D(X)=\sigma^2$$

证明

$$\begin{aligned}
&ES^2=E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\right]=\frac{1}{n-1}E\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\\
&=\frac{1}{n-1}E\sum_{i=1}^{n}\left(X_i^2-2\bar{X}X_i+\bar{X}^2\right)=\frac{1}{n-1}E\left[\sum_{i=1}^{n}X_i^2-2\bar{X}\sum_{i=1}^{n}X_i+\sum_{i=1}^{n}\bar{X}^2\right]\\
&=\frac{1}{n-1}E\left[\sum_{i=1}^{n}X_i^2-2\bar{X}\cdot n\bar{X}+n\bar{X}^2\right]\\
&=\frac{1}{n-1}\left[\sum_{i=1}^{n}EX_i^2-nE(\bar{X})^2\right]\\
&=\frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\mu^2+\sigma^2\right)-n\left(\mu^2+\frac{1}{n}\sigma^2\right)\right]=\sigma^2
\end{aligned}$$

定义 6.7 (分位数) 若 $P(X>\mu_\alpha)=\alpha$, 则称 μ_α 为标准正态分布的上 α 分位数。若 $P(|X|>\mu_{\frac{\alpha}{2}})=\alpha$, 则称 $\mu_{\frac{\alpha}{2}}$ 为标准正态分布的双侧 α 分位数。

6.1 卡方分布

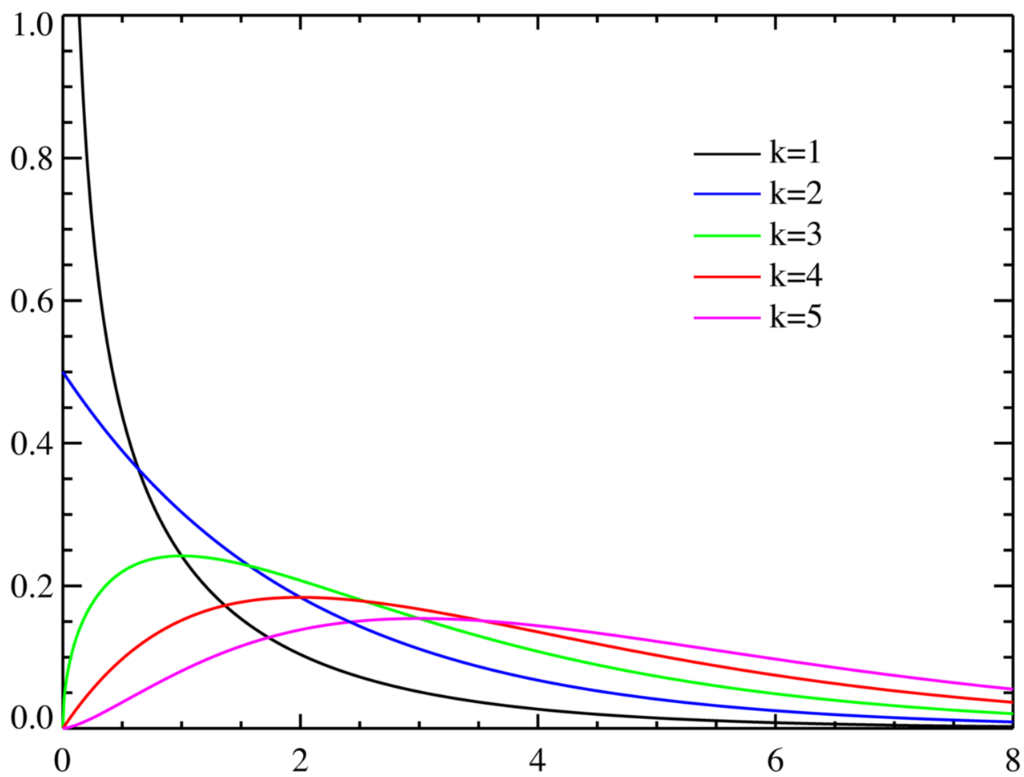
定义 6.8 (卡方分布) 设 $X_1,X_2,\ldots,X_n$ 相互独立，且都服从标准正态分布 $N(0,1)$ 则

$$\sum_{i=1}^{n}X_i^2\sim\chi^2(n)=\Gamma(1/2,n/2)$$

一般的，自由度为 n 的 $\chi^2(n)$ 的密度函数为

$$f(x)=\begin{cases}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}e^{-\frac{x}{2}}x^{\frac{n}{2}-1}, x>0\\ 0,\quad x\leq 0\end{cases}$$

其中 $\Gamma(\alpha)=\int_0^{+\infty}t^{\alpha-1}e^{-t}dt$ 概率密度函数的图像如下图所示：



卡方分布的性质

- $E\left(\chi^2(n)\right)=n, D\left(\chi^2(n)\right)=2n$

  设 $\chi^2(n)=\sum_{i=1}^n X_i^2, X_i \sim N(0,1)$,$X_1, X_2 \ldots, X_n$ 相互独立，则

$$
\begin{aligned}
& E\left(X_i\right)=0, D\left(X_i\right)=1, E\left(X_i^2\right)=1 \\
& E\left(\chi^2(n)\right)=E\left(\sum_{i=1}^n X_i^2\right)=n \\
& E\left(X_i^4\right)=\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^4 e^{-\frac{x^2}{2}} \mathrm{d}x=3 \\
& D\left(X_i^2\right)=E\left(X_i^4\right)-E^2\left(X_i^2\right)=2 \\
& D\left(\chi^2(n)\right)=D\left(\sum_{i=1}^n X_i^2\right)=2n
\end{aligned}
$$

- 若 $X_1=\chi^2(n_1), X_2=\chi^2(n_2), X_1, X_2,$ 相互独立，则 $X_1+X_2=\chi^2(n_1+n_2)$
- 当 $n \to \infty$，则 $\chi^2(n)$ 近似于正态分布。

定理 6.2 (核心定理) $X_1, \cdots, X_n \sim N\left(\mu, \sigma^2\right)$ 且独立同分布，则

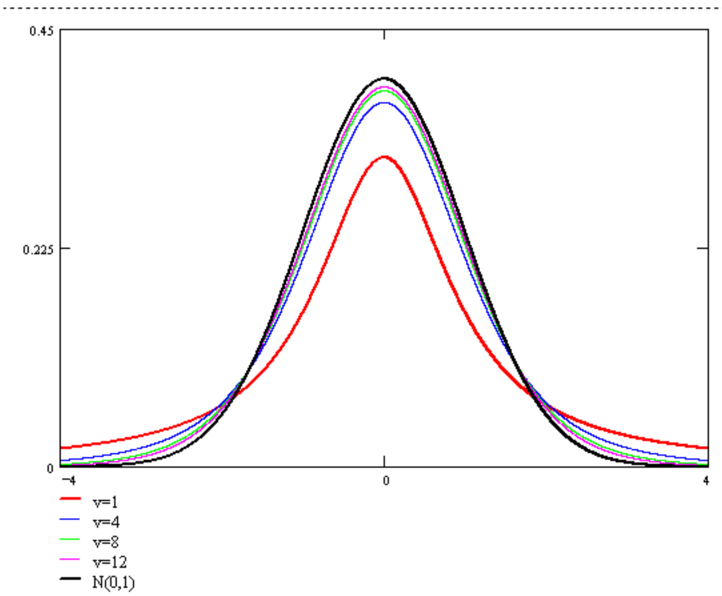
- $\overline{X}$ 与 S^2 相互独立
- $\chi^2=\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

6.2 t 分布

设 $X \sim N(0,1), Y \sim \chi^2(n)$,X，Y 相互独立，$T = \frac{X}{\sqrt{\frac{Y}{n}}}$, 则称 T 服从自由度为 n 的 T 分布，记为 $T \sim t(n)$ 其密度函数为

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \qquad -\infty < t < \infty$$

图像如下



性质

- 具有自由度 n 的 t 分布 $t \sim t(n)$，其数学期望与方差为 $E(t) = 0, D(t) = \frac{n}{n-2}(n > 2)$
- t 分布的密度函数关于 t=0 对称，当能充分大的时候，密度函数近似于标准正态分布函数
-

定理 6.3 $X_1, \cdots, X_n \sim N\left(\mu, \sigma^2\right)$ 且独立同分布，则

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1)$$

证明

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$$

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$$

又由于 $\frac{(n-1)S^2}{\sigma^2}$ 与 $\overline{X}$ 相互独立，所以

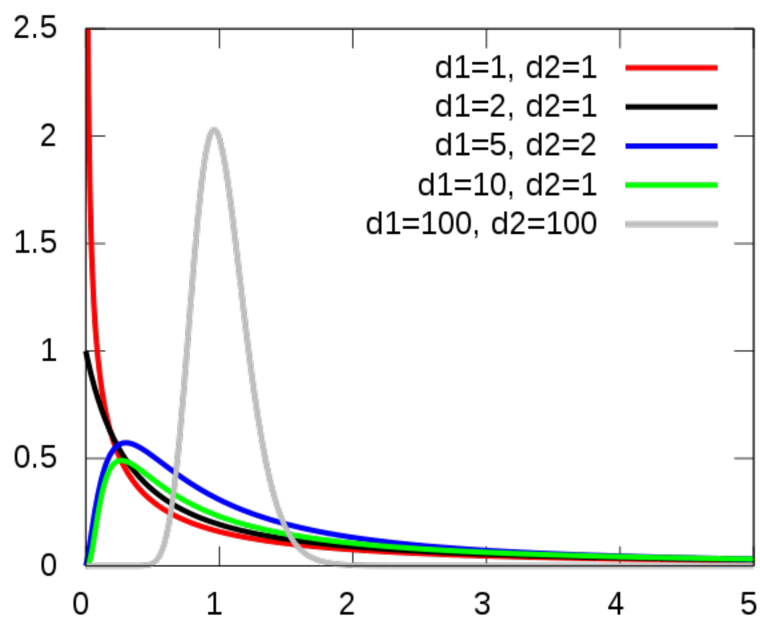
$$\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \Big/ \frac{S}{\sigma} = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t(n-1)$$

6.3 F 分布

定义 6.9 设 $X \sim \chi^2(n)$,X 和 Y 相互独立，令 $F = \frac{X/n}{Y/m}$ 则称 F 服从第一自由度为 n，第二自由度为 m 的 F 分布。其密度函数为

$$f(t,n,m) = \left(\frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)}\left(\frac{n}{m}\right)^{\frac{n}{2}} t^{\frac{n}{2}-1}\left(1+\frac{n}{m}t\right)^{-\frac{n+m}{2}} t>0\right).$$

图像为:



定理 6.4 设 $X_1, X_2, \ldots, X_n$ 与 $Y_1, Y_2 \ldots, Y_m$ 分别是来自正态总体 $X \sim N(\mu_1, \sigma_1^2)$ 与 $Y \sim N(\mu_2, \sigma_2^2)$

的相互独立的简单随机样本，令

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \bar{Y} = \frac{1}{m}\sum_{j=1}^{m} Y_j$$
$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 \quad S_2^2 = \frac{1}{m-1}\sum_{j=1}^{m} \left(Y_j - \bar{Y}\right)^2$$

则

$$\frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2(n-1) \quad \frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2(m-1)$$

所以

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F(n-1, m-1)$$

若 $\sigma_1 = \sigma_2$ 则

$$\frac{S_1^2}{S_2^2} \sim F(n-1, m-1)$$

性质

- F 分布的数学期望为 $E(F) = \frac{n_2}{n_2-2}$
- 若 $F \sim F(n,m)$, 则 $\frac{1}{F} \sim F(m,n)$

例题 21

从正态总体 $X \sim N(\mu, \sigma^2)$ 中抽取了 n=20 的样本 $(X_1, X_2 \ldots, X_{20})$

- 求 $P(0.27\sigma^2 \leq \frac{1}{20}\sum_{i=1}^{2} 0(X_i - \overline{X})^2 \leq 1.76\sigma^2)$
- 求 $P(0.27\sigma^2 \leq \frac{1}{20}\sum_{i=1}^{2} 0(X_i - \mu)^2 \leq 1.76\sigma^2)$

解答

(1) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

$\frac{19S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\bar{X})^2 \sim \chi^2(19)$

$P\left(0.37\sigma^2 \leq \frac{1}{20}\sum_{i=1}^{20}(X_i-\bar{X})^2 \leq 1.76\sigma^2\right)$

$= P\left(7.4 \leq \frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\bar{X})^2 \leq 35.2\right)$

$= P\left(\frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\bar{X})^2 \geq 7.4\right) - P\left(\frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\bar{X})^2 \geq 35.2\right)$

$= 0.99 - 0.01 = 0.98$

(2) $\sum_{i=1}^{20}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(20)$

$P\left(0.37\sigma^2 \leq \frac{1}{20}\sum_{i=1}^{20}(X_i-\mu)^2 \leq 1.76\sigma^2\right)$

$= P\left(7.4 \leq \sum_{i=1}^{20}\left(\frac{X_i-\mu}{\sigma}\right)^2 \leq 35.2\right)$

$= P\left(\frac{20}{i-1}\left(\frac{X_i-\mu}{\sigma}\right)^2 \geq 7.4\right) - P\left(\sum_{i=1}^{20}\left(\frac{X_i-\mu}{\sigma}\right)^2 \geq 35.2\right)$

$= 0.995 - 0.025 = 0.97$

7 参数估计

定义 7.1 (参数估计) 设有一个统计总体, 总体的分布函数为 $F(x,\theta)$ ，其中 θ 为未知参数 (θ 可以是向量). 现从该总体抽样，得样本

$$X_1, X_2, \ldots, X_n$$

要依据该样本对参数 θ 作出估计, 或估计 θ 的某个已知函数 $g(\theta)$, 这类问题称为参数估计

定义 7.2 (最大似然法) 是在总体分布已知条件下使用的一种参数估计方法.

定义 7.3 (最大似然估计法) 设 $x_1, \ldots, x_n$ 为相应于样本 $X_1, \ldots, X_n$ 的一个样本值,则样本 $X_1, X_2, \ldots, X_n$ 取到观察值 $x_1, x_2 \ldots, x_n$ 的概率为

$$L(\theta) = L(x_1, x_2, \ldots, x_n; \theta) = \Pi p(x_i; \theta)$$

$L(\theta)$ 称为样本的似然函数。