1. 随机事件及其概率

1.1. 随机试验和随机事件

1.1.1. 随机试验

概率论中将满足下面三个条件的试验称为随机试验，简称试验：

(1) 可在相同的条件下重复进行；

(2) 所有的结果是明确可知的：

(3) 试验之前不能确定哪一个结果会发生.

1.1.2. 样本空间

随机试验的所有可能结果所组成的樂合称为样本空间，常记为Ω，Ω中的元素称为样本点.

1.1.3. 随机事件

样本空间的子集，即试验的结果称为随机事件，简称事件，由一个样本点組成的单点集，称为基本事件.

另外，两个特殊的事件为：

必然事件Ω----每次试验中一定发生的事件；

不可能事件$\varnothing$----每次试验中一定不发生的事件.

1.2. 事件的关系、运算及其运算律

1.2.1. 事件 A 与 B 之和（并）

$A \cup B$（或 $A+B$）表示事件 A 与 B 至少有一个发生

1.2.2. 事件 A 与 B 的积

$A \cap B$（或 AB）表示事件 A 与 B 同时发生

1.2.3. 事件 A 与 B 的差

$A-B$ 表示事件A发生而B不发生

性质：$A-B=A\overline{B}$ 或 $A \cap \overline{B}$.

1.2.4. 事件的包含

若时间A发生必然导致事件B发生，则称事件B包含A（或A包含于B），记为$B \supset A$.

1.2.5. 事件相等

若$A \supset B$ 且 $B \subset A$，则称事件 A 与 B 相等，记为 $A = B$.

1.2.6. 互斥事件

在试验中，若事件A与B不能同时发生，即$A \cap B = \varnothing$，则称A、B为互斥事件.

1.2.7. 对立事件

每次试验中，“事件A不发生”的事件称为A的对立事件或逆事件，A的对立事件记为$\overline{A}$.
性质：

(1) $A + \overline{A} = \Omega$（必然事件）

(2) $A\overline{A} = \varnothing$（不可能事件）

1.2.8. 事件的运算律

(1) 吸收律 若$A \subset B$，则$A \cup B = B$，$AB = A$，$A\overline{B} = \varnothing$.
(2) 交换律 $A \cup B = B \cup A$，$AB = BA$.
(3) 结合律 $(A \cup B) \cup C = A \cup (B \cup C)$，$(A \cap B) \cap C = A \cap (B \cap C)$.
(4) 分配律 $(A \cup B)C = (AC) \cup (BC)$，$A \cup (BC) = (A \cup B)(A \cup C)$.
(5) 德摩根律[1] $\overline{A_1 \cup A_2} = \overline{A_1} \cap \overline{A_2}$，$\overline{A_1 \cap A_2} = \overline{A_1} \cup \overline{A_2}$.

1.3. 概率及其性质

1.3.1. 概率的定义

设随机试验E的样本空间为Ω，则称满足下列条件的事件集上的函数$P(\cdot)$为概率：
(1) 对于任意事件A，$P(A) \geqslant 0$；
(2) 对于必然事件Ω，$P(\Omega) = 1$；
(3) 设 $A_1, A_2, \cdots, A_n, \cdots$ 为两两互不相容的事件，即$A_iA_j = \varnothing(i \neq j, i, j = 1, 2, \cdots)$，则

$$P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k).$$

1.3.2. 概率的性质

(1) $P(\overline{A}) = 1 - P(A)$；
(2) $P(A - B) = P(A) - P(AB)$，特别，当$B \subset A$时，$P(A - B) = P(A) - P(B)$，且$P(B) \leqslant P(A)$；
(3) $P(A \cup B) = P(A) + P(B) - P(AB)$,

[1] 长杆变短杆，开口换方向

$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC)$ ；
特别的，若A，B互斥，则$P(A\cup B)=P(A)+P(B)-P(AB), P(A-B)=P(A)$

1.4. 三大概型

1.4.1. 古典概型

如果随机试验E满足下列条件:

(1) 试验的样本空间Ω的元素只有有限个，

(2) 样本空间中每个元素，即基本事件发生的可能性相同，则称此试验为古典概型，对于古典概型，事件A的概率有下列计算公式：

$$P(A)=\frac{A\text{中基本事件数}}{\Omega\text{中基本事件总数}}$$

1.4.2. 几何概型

如果随机试验E的样本空间Ω为欧式空间中的一个区域，且每个样本点的出现具有等可能性，则称此试验为几何概型，对于几何概型，事件A的概率有下列计算公式：

$$P(A)=\frac{A\text{的度量(长度、面积、体积)}}{\Omega\text{的度量(长度、面积、体积)}}.$$

1.4.3. 伯努利(Bernoulli)模型

如果试验E的结果只有两个，A与$\overline{A}$，则称此试验为伯努利概型（试验），若将伯努利试验独立重复n次，则称为n重伯努利概型，简称伯努利概型，在伯努利概型中，若$P(A)=p$，则n次试验中事件A发生k次的概率为

$$P_n(k)=C_n^k p^k(1-p)^{n-k}, k=0,1,\ \cdots, n.$$

1.5. 三大概率公式

1.5.1. 条件概率与乘法公式

设A，B是两个事件，且$P(A)>0$，称

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

为在事件A发生条件下事件B发生的条件概率.

乘法公式：设$P(A_1)>0, P(A_2)>0$，则

$$P(A_1A_2)=P(A_1)P(A_2|A_1)=P(A_2)P(A_1|A_2).$$

一般地，设$P(A_1A_2\cdots A_{n-1})>0$，则

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1\cdots A_{n-1})$$

1.5.2. 全概率公式

设$B_1,B_2,\cdots,B_n$为一完备事件组，即$B_iB_j=\varnothing,i\neq j$，$\bigcup_{i=1}^{n}B_i=\Omega,P(B_i>0),i=1,2,\cdots,n$,则对事件A有

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

1.5.3. 贝叶斯(Bayes)公式

设$B_1,B_2,\cdots,B_n$为一完备事件组，$P(B_i)>0,i=1,2,\cdots,n$，$P(A)>0$，则有

$$P(B_j|A)=\frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n}P(A|B_i)P(B_i)},j=1,2,\cdots,n.$$

1.6. 事件的独立性

1.6.1. 两个事件的独立性

设A,B是两个事件，若有等式

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

则称A与B相互独立.

1.6.2. 两事件相互独立的充要条件

① $P(AB)=P(A)P(B)$

② $P(B|A)=P(B)\quad(P(A)>0)$；或 $P(A|B)=P(A)\quad(P(B)>0)$

③ 四对事件 A与B，A与$\overline{B}$，$\overline{A}$与B，$\overline{A}$与$\overline{B}$中有一对相互独立，则另外三对也相互独立.

1.6.3. n个事件的两两独立于相互独立 （$n \geqslant 3$）

(1) n个事件$A_1, A_2, \cdots, A_n$两两独立：其中任何两个时间相互独立，即有下列$C_n^2 = \dfrac{n(n-1)}{2}$个等式成立：
$P(A_iA_j) = P(A_i)P(A_j)$ （$1 \leqslant i < j \leqslant n$）

(2) n个事件$A_1, A_2, \cdots, A_n$相互独立，其中任意k个不同事件都相互独立，即有下列$2^n - n - 1$个等式成立；
$P(A_{i1}A_{i2} \cdots A_{ik}) = P(A_{i1})P(A_{i2}) \cdots P(A_{ik})$，其中$i \leqslant i_1 < i_2 < \cdots < i_k \leqslant n, k = 2, 3, \cdots, n$,

(3) n个事件两两独立于相互独立的关系:
$A_1, A_2, \cdots, A_n$相互独立 $\Rightarrow A_1, A_2, \cdots, A_n$两两独立
$A_1, A_2, \cdots, A_n$两两独立 $\nRightarrow A_1, A_2, \cdots, A_n$相互独立

(4) n个事件相互独立的结论：若$A_1, A_2, \cdots, A_n$相互独立，则由其中任意部分事件所产生的事件与另一部分事件所产生的事件相互独立.

2. 随机变量及其分布

2.1. 分布函数

(1) 定义　设X为随机变量，x为任意实数，则函数 $F(x) = P(X \leqslant x)$称为随机变量X的分布函数
(2) 性质：
① $0 \leqslant F(x) \leqslant 1, -\infty < x < +\infty$；
② $F(x)$是单调不减的函数，即$x_1 < x_2$时，有$F(x_1) \leqslant F(x_2)$;
③ $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to +\infty} F(x) = 1$；
④ $\lim\limits_{x \to x_0^+} F(x) = F(x_0)$，即$F(x)$是右连续的；

2.2. 离散型随机变量

2.2.1. 分布律

设X的所有取值为$x_k(k = 1, 2, \cdots)$，则称事件$\{X = x_k\}$的概率即$P\{X = x_k\} = p_k$，或下列表格.

X	x_1	x_2	$, \cdots ,$	x_n	$\cdots$
P	p_1	p_2	$, \cdots ,$	p_n	$\cdots$

为X的分布律

2.2.2. 分布律的性质

(1) $p_i \geqslant 0, i = 1, 2, \cdots, n, \cdots$

(2) $\sum_{k=1}^{\infty} p_k = 1$

2.2.3. 分布函数$F(x)$

设X的分布律为$P\{X = x_k\} = p_k$，$k = 1, 2, \cdots$，则X的分布函数为$F(x) = P(X \leqslant x) = \sum_{x_k \leqslant x} P\{X = x_k\}$.

若已知X的分布函数$F(x)$，则可求得X的分布律:

$$P\{X = x_k\} = F(x_k) - \lim_{x \to x_k^-} F(x), k = 1, 2, \cdots$$

2.3. 连续型随机变量

2.3.1. 定义

若随机变量X的分布函数$F(x)$，可以表示成非可积负函数$f(x)$的积分形式:

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

则称X为连续型随机变量，$f(x)$称为X的概率密度函数，简称概率密度。

2.3.2. 概率密度函数的性质

(1) $f(x) \geqslant 0$

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

2.3.3. 连续型随机变量分布函数的性质

(1) $F(x)$是关于x的连续函数，对于任何实数c，有$P\{X = c\} = 0$

(2) 若$f(x)$在x处连续，则有$F'(x) = f(x)$

$P(x_1 \leqslant X \leqslant x_2) = P(x_1 < X \leqslant x_2) = P(x_1 \leqslant X < x_2) = P(x_1 < X < x_2)$

$= F(x_2) - F(x_1) = \int_{x1}^{x_2} f(t)dt$

2.4. 离散型分布（五大分布）

2.4.1. $0-1$分布

设一次伯努利试验中事件A发生X次，则X服从$0-1$分布，其分布律为

X	1	0
P	p	$1-p$

其中p为时间A出现的概率，$0<p<1$.

2.4.2. 二项分布

设n重伯努利试验中事件A发生X次，则X服从二项分布，记作$X\sim B(n,p)$，其分布律为

$$P(X=k)=C_n^k p^k q^{n-k},\ k=0,1,2,\ \cdots,n,$$

2.4.3. 几何分布

进行独立重复试验，设每次试验成功的概率为p，失败的概率为$q=1-p(0<p<1)$，将试验进行到出现一次成功为止，以X表示所需的试验次数，则X服从参数为p的几何分布，记为$X\sim G(p)$，其分布律为

$$P(X=k)=q^{k-1}p,\ k=1,2,3,\ \cdots,\ \text{其中}\ p\geqslant 0,q=1-p$$

2.4.4. 超几何分布

N件产品，其中M件次品，从中任取n件，有X个次品，则随机变量X服从参数为n,N,M的超几何分布，通常记为$X\sim H(n,N,M)$，其分布律为

$$P(X=i)=\frac{C_M^i C_{N-M}^{n-i}}{C_N^n},0\leqslant i\leqslant n\leqslant N,i\leqslant M$$

2.4.5. 泊松分布

设随机变量X的分布律为

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

则称随机变量X服从参数为λ的泊松分布，记为$X\sim P(\lambda)$.

2.5. 连续型分布（三大分布）

2.5.1. 均匀分布

设随机变量X的值等可能地落在$[a,b]$内，则称X在$[a,b]$上服从均匀分布，记为$X\sim U(x,y)$，其密度分布为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a\leqslant x\leqslant b,\\ 0, & others.\end{cases}$$

分布函数为

$$f(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\leqslant x<b\\ 1 & x\geqslant b\end{cases}$$

2.5.2. 指数分布

若连续型随机变量X的概率密度为

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

其中$\lambda>0$，则称X服从参数为λ的指数分布，记为$X\sim E(\lambda)$，其分布函数

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

2.5.3. 正态分布

(1) 一般正态分布

(a) 定义 若随机变量X的密度函数为

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

其中$\sigma>0$，μ与σ均为常数，则称X服从参数为μ和σ的正态分布，记为$X\sim N\left(\mu,\sigma^2\right)$.其分布函数

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

(b) $f(x)$和$F(x)$的性质如下：

(i) $f(x)$的图形是关于$x=\mu$对称，即$f(\mu+x)=f(\mu-x)$；
(ii) $F(\mu+x)+F(\mu-x)=1$；
(iii) $F(\mu)=\frac{1}{2}$；
(iv) 当$x=\mu$时，$f(\mu)=\frac{1}{\sqrt{2\pi\sigma}}$为最大值；
(v) $f(x)$以x轴为渐近线.

(2) 标准正态分布
(a) 定义 当$u=0$，$\sigma^2=1$，即X的密度函数为

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

则称X服从标准正态分布，记为$X\sim N(0,1)$，其分布函数为

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x.$$

(b) $\varphi(x)$和$\Phi(x)$的性质如下：
(i) $\varphi(-x)=\varphi(x)$
(ii) $\Phi(-a)+\Phi(a)=1$
(iii) $\Phi(0)=\frac{1}{2}$

(3) 一般正态分布与标准正态分布之间的关系：如果$X\sim N\left(\mu,\sigma^2\right)$，则$\frac{X-\mu}{\sigma}\sim N(0,1)$,

$$P(x_1<X\leqslant x_2)=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right).$$

2.6. 随机变量函数的分布

随机变量Y是随机变量X的函数$Y=g(X)$，其中$g(x)$为连续函数或分段函数，现要求$Y=g(X)$的概率分布，分两种情形讨论：

2.6.1. X是离散型随机变量

已知X的分布列为

X	x_1	x_2	, $\cdots$,	x_n	$\cdots$
P	p_1	p_2	, $\cdots$,	p_n	$\cdots$

显然，$Y=g(x)$的取值只可能是$g(x_1),g(x_2),\cdots,g(x_n),\cdots$，若$g(x_i)$互不相等，则$Y$的分布列如下：

Y	$g(x_1)$	$g(x_2)$	$,\cdots,$	$g(x_n)$	$\cdots$
P	p_1	p_2	$,\cdots,$	p_n	$\cdots$

若有某些$g(x_i)$相等，则应将对应的p_i相加作为$g(x_i)$的概率

2.6.2. X是连续型随机变量

分布函数法：先按分布函数的定义求得Y的分布函数$F_Y(y)$，再通过求导得到Y的概率密度$f_Y(y)$，即

$$F_Y(y) = P(Y \leqslant y) = P(g(x) \leqslant y) = \int_{g(x)\leqslant y} f_X(x)\mathrm{d}x，\ f_Y(y) = \frac{\mathrm{d}F_Y(y)}{dy};$$

3. 多维随机变量及其分布

3.1. 二维随机变量联合分布函数及其性质

3.1.1. 二维RV联合分布函数定义

设(X, Y)是二维随机变量，对任意实数x和y，称

$$F(x, y) = P\{X \leqslant x, Y \leqslant y\}$$

为(X, Y)的分布函数，又称联合分布函数.

3.1.2. 二维RV联合分布函数性质

(1) $0 \leqslant F(x, y) \leqslant 1$;
(2) $F(x, y)$对x和y都是单调非减的，即对任意的x_1,x_2，若$x_1 > x_2$，则有$F(x_1, y) \geqslant F(x_2, y)$;
(3) $\lim\limits_{\substack{x\to+\infty \\ y\to+\infty}} F(x, y) = 1$，$\lim\limits_{\substack{x\to-\infty \\ y\to-\infty}} F(x, y) = \lim\limits_{x\to-\infty} F(x, y) = \lim\limits_{y\to-\infty} F(x, y) = 0$;
(4) $F(x, y)$对x和y都是右连续.

3.1.3. 联合分布函数的几何意义

$F(x, y)$在(x, y)的函数值就是随机点(X, Y)在$X = x$左侧和$Y = y$下方的无穷矩形内的概率，对有限矩形域有：

$$P\{x_1 < X < x_2, y_1 < Y < y_2\} = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1).$$

3.1.4. 边缘分布函数

在试验涉及二维随机变量(X, Y)的前提下，单个随机变型X与Y的分布函数$F_X(x)$和$F_Y(y)$称为二维随机变量(X, Y)的关于X和关于Y的边缘分布函数，即

$$F_X(x) = P\{X \leqslant x\} = \lim_{y \to +\infty} F(x, y)$$
$$F_Y(y) = P\{Y \leqslant y\} = \lim_{x \to +\infty} F(x, y)$$

3.2. 二维离散型随机变量的联合概率分布、边缘概率分布及条件概率分布

3.2.1. 联合分布律

(1) 定义

设离散型随机变量(X, Y)的所有可能取值为$(x_i, y_i)(i, j = 1, 2, \cdots)$且事件$\{(X, Y) = (x_i, y_j)\}$的概率为$p_{ij}$，称

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

为(X, Y)的联合分布律，联合分布律有时也用下面的概率分布表来表示：

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	$P\{X = x_i\}$
x_1	P_{11}	P_{12}	$\cdots$	P_{1j}	$\cdots$	$P\{X = x_1\}$
x_2	P_{21}	P_{22}	$\cdots$	P_{2j}	$\cdots$	$P\{X = x_2\}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
x_i	P_{i1}	P_{i2}	$\cdots$	P_{ij}	$\cdots$	$P\{X = x_i\}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
$P\{Y = y_j\}$	$P\{Y = y_1\}$	$P\{Y = y_2\}$	$\cdots$	$P\{Y = y_j\}$	$\cdots$	1

(2) 性质

(a) $P_{ij} \geqslant 0(i, j = 1, 2, \cdots)$；

(b) $\sum_i \sum_j p_{ij} = 1.$

3.2.2. 边缘分布律

设随机变盘(X, Y)的联合分布律为

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

则X的边缘分布为

$$P(X=x_i)=\sum_j p_{ij}(i,j=1,2,\ \cdots);$$

Y的边缘分布为

$$P(Y=y_j)=\sum_i p_{ij}(i,j=1,2,\ \cdots).$$

3.2.3. 条件分布律

二维离散型随机变型(X,Y)的联合分布律为

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

在已知$X=x_i$的条件下，Y取值的条件分布律为

$$P(Y=y_j|X=x_i)=\frac{P\{X=x_i,Y=y_j\}}{P(X=x_i)},$$

在已知$Y=y_j$的条件下，X取值的条件分布律为

$$P(Y=y_i|X=x_j)=\frac{P\{X=x_i,Y=y_j\}}{P(Y=y_j)},$$

其中$P(X=x_i)$，$P(Y=y_j)$分别为X，Y的边缘分布律.

3.3. 二维连续型随机变量的联合概率密度、边缘概率密度及条件概率密度

3.3.1. 连续型随机变量

设$F(x,y)$为二维随机变量(X,Y)的分布函数，如果存在非负函数$f(x,y)$使得对任意实数x，y有
$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v$，则称$(X,Y)$为二维连续型随机变量，$f(x,y)$为$(X,Y)$的联合概率密度.

3.3.2. 联合概率密度$f(x,y)$的性质：

(1) $f(x,y) \geqslant 0$；

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x\mathrm{d}y = 1$.

3.3.3. 边缘概率密度

当(X,Y)为连续型随机变量，并且其联合概率密度为$f(x,y)$，则X和Y的边缘概率密度为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y,\ f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x.$$

3.3.4. 条件概率密度

二维连续型随机变量(X,Y)的联合概率密度为$f(x,y)$，则在已知$Y=y$的条件下，X的条件概率密度为

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

在已知$X=x$的条件下，Y的条件概率密度为

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

其中$f_x(x) > 0, f_Y(y) > 0$分别为X，Y的边缘概率密度.

3.4. 常见的二维分布

3.4.1. 二维均匀分布

设随机变量(X,Y)的概率密度为

$$f(x,y) = \begin{cases} 1/S_D, & (x,y) \in D \\ 0, & others \end{cases}$$

其中S_D为区城D的面积，则称(X,Y)服从D上的均匀分布，记为$(X,Y)\sim U(D)$.

3.4.2. 二维正态分布

设随机变量(X,Y)的概率密度为

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

其中$\mu_1,\mu_2,\sigma_1>0,\sigma_2>0,|\rho|\leqslant 1$，则称$(X,Y)$服从二维正态分布，记为$(X,Y)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$

3.5. 随机变量的独立性

3.5.1. 一般型随机变量

相互独立：设(X,Y)的分布函数为$F(x,y)$，边缘分布函数为$F_X(x)$和$F_Y(y)$，如果对一切x,y有$F(x,y)=F_X(x)F_Y(y)$则称X与Y相互独立.

3.5.2. 离散型随机变量

X与Y相互独立的充要条件是：$P(X=x_i,Y=y_j)=P(X\ =x_i)P(Y=y_j)$

3.5.3. 连续型随机变量

X与Y是相互独立的充要条件是$f(x,y)=f_X(x)f_Y(y)$

3.5.4. 随机变量函数的独立性

若X与Y独立，h,g为X与Y连续函数，则：$h(X)$和$g(Y)$独立.

3.5.5. 二维正态分布关于独立的相关结论

1. 若$(X,Y)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$，则$X\sim N\left(\mu_1,\sigma_1^2\right),Y\sim N\left(\mu_2,\sigma_2^2\right)$；但是若$X\sim N\left(\mu_1,\sigma_1^2\right)$，$Y-N\left(\mu_2,\sigma_2^2\right)$，则$(X,Y)$未必是二维正态分布.
2. 若$(X,Y)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$，则随机变量$X,Y$的非零线性组合$Z=aX+bY\neq 0$服从一维正态分布.
3. 若$X\sim N\left(\mu_1,\sigma_1^2\right)$，$Y\sim\left(\mu_2,\sigma_2^2\right)$，且$(X,Y)$相互独立，则$(X,Y)$服从二维正态分布.
4. 若$(X,Y)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$，则$X$和$Y$相互独立的充要条件是$\rho=0$
5. 若$X_1,X_2,\cdots,X_n$相互独立且服从正态分布，则$X_1,X_2,\cdots,X_n$的线性组合服从正态分布

3.6. 两个随机变量函数的分布

3.6.1. 两个随机变量函数的定义

设X,Y为两个随机变量，$z=g(x,y)$为二元连续函数，则称$Z=g(X,Y)$为随机变量X,Y的函数，显然两随机变量函数

是一维随机变量.

3.6.2. 二维离散型随机变量函数的分布律

$Z=g(X,Y)$也是离散型随机变量，其分布律为

$$P\{Z=g(x_i,y_j)\}=P\{X=x_i,Y=y_j\}=P_{ij},$$

如果有若干个$g(x_i,y_j)$相同，则合并诸项为一项，并将相应的概率相加作为Z取值为$g(x_i,y_j)$的概率.

Z的分布函数为：$F_Z(z)=P\{Z\leqslant z\}=P\{g(X,Y)\leqslant z\}=\sum\limits_{g(x_i,y_j)\leqslant z}P(X=x_i,Y=y_j)$.

3.6.3. 二维连续型随机变量函数的分布

设(X,Y)的联合概率密度为$f(x,y)$，则$Z=g(X,Y)$的分布函数为

$$F_Z(z)=P\{Z\leqslant z\}=P\{g(X,Y)\leqslant z\}=\iint\limits_{g(x,y)\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y$$

3.6.4. 混合型

(X和Y中有一个是离散型随机变量，另一个为连续型随机变量）利用分布函数法求之.

3.7. 极值分布

设X与Y是两个相互独立的随机变量，其分布函数分别为$F_X(x)$和$F_Y(y)$，则
$Z=\max(X,Y)$的分布函数为：

$$F_{\max}(z)=F_X(z)F_Y(z);$$

$Z=\min(X,Y)$的分布函数为：

$$F_{\min}(z)=1-[1-F_X(z)][1-F_Y(z)].$$

特别地当$X_1,X_2,\cdots,X_n$相互独立时
$Z=\max(X_1,X_2,\cdots,X_n)$的分布函数为：

$$F_{\max}(z)=F_{X_1}(z)F_{x_2}(z)\cdots F_{x_n}(z)$$

$Z=\min(X_1,X_2,\cdots,X_n)$的分布函数分别为：

$$F_{\min}(z)=1-[1-F_{X_1}(z)][1-F_{x_2}(z)]\cdots[1-F_{x_n}(z)]$$

4. 随机变量的数字特征

4.1. 随机变量的数学期望

4.1.1. 一维随机变量的数学期望

(1) 离散型随机变量的数学期望

设离散型随机变量X的分布律为$P\{X=X_i\}=p_i(i=1,2,\cdots)$.若无穷级数$\sum_{i=1}^{\infty}x_ip_i$绝对收敛，则称无穷级数$\sum_{i=1}^{\infty}x_ip_i$的和为随机变量$X$的数学期望，记作$E(X)$，即

$$E(X)=\sum_{i=1}^{\infty}x_ip_i$$

(2) 连续型随机变量的数学期望

设X为连续型随机变量，其概率密度为$f(x)$，若反常积分$\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$绝对收敛，则称反常积分$\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$的值为随机变量$X$的数学期望，记为$E(X)$，即

$$E(x)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$$

4.1.2. 随机变量函数的数字期望

(1) 一维随机变量函数的数学期望

设Y是随机变量X的函数；$Y=g(X)$（g是连续函数或分段连续函数).

设离散型随机变量X具有分布律$p\{X=x_i\}=p_i$，$i=1,2,\cdots$，且无穷级数$\sum_{i=1}^{\infty}g(x_i)p_i$绝对收敛；则有：

$$E(Y)=E[g(X)]=\sum_{i=1}^{\infty}g(x_i)p_i$$

设连续型随机变量X具有概率密度函数$f(x)$，且反常积分$\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$绝对收敛，则有：

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

(2) 二维随机变量函数的数学期望

设Z是随机变量X, Y的函数：$Z = g(X, Y)$（g是连续函数），那么，Z是一个一维随机变量.

设二维离散型随机变量(X, Y)具有分布律$P\{X = x_i, Y = y_i\} = p_{ij}(i, j = 1, 2, \cdots)$；则有

$$E(Z) = E[g(X, Y)] = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} g(x, y)P_{ij}$$

或者，设Z具有分布律$P\{Z = z_i\} = p_i(i = 1, 2, \cdots)$，则有：

$$E(Z) = \sum_{i=1}^{\infty} z_i p_i$$

这里的无穷级数绝对收敛.

设二维连续型随机变量(X, Y)具有联合概率密度$f(x, y)$，则有

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

设Z具有概率密度$f_Z(z)$，则有：

$$E(Z) = \int_{-\infty}^{+\infty} z f_Z(z)\mathrm{d}z$$

这里的反常积分绝对收敛.

4.1.3. 数学期望的性质

(1) $E(C) = C$，（C为常数）

(2) $E(CX) = CE(X)$；

(3) $E(X+C) = E(X)+C$；

(4) $E(X+Y) = E(X)+E(Y)$；

综合(1)(2)(4)有$E(\sum_{i=1}^{n} C_i X_i) = \sum_{i=1}^{n} C_i E(X_i)$；

4.2. 随机变量的方差

4.2.1. 方差及标准差的概念

设X是一个随机变量，若$E\left\{[X-E(X)]^2\right\}$存在，则称$E\left\{[X-E(X)]^2\right\}$为$X$的方差，记为$D(X)$，即

$D(X) = E\left\{[X - E(X)]^2\right\}$

方差的算术平方根$\sqrt{D(X)}$称为X的均方差或标准差.

(1) 离散型随机变量的方差

$$D(X) = \sum_k [x_k - E(X)]^2 p_k$$

(2) 连续型随机变量的方差

$$D(X)=\int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)\mathrm{d}x$$

4.2.2. 方差的计算公式

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

4.2.3. 方差的性质

(1) $D(C) = 0$, （C为常数）；
(2) $D(CX) = C^2 D(X)$;
(3) $D(X + C) = D(X)$;
(4) 若$X_1, X_2, \cdots, X_m$两两独立或两两不相关，则
$D(C_1X_1 + C_2X_2 + \cdots + C_mX_m) = C_1^2D(X_1) + C_2^2D(X_2) + \cdots + C_m^2D(X_m)$.

分布	记号	期望	方差
$0-1$分布	$B(1, p)$	p	$p(1-p)$
二项分布	$B(n, p)$	np	$np(1-p)$
泊松分布	$P(\lambda)$	λ	λ
几何分布	$G(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
超几何分布	$H(n, M, N)$	$\frac{nM}{N}$	$\frac{nM}{N}\left(1-\frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$

均匀分布	$U(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布	$E(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
正态分布	$N\left(\mu,\sigma^2\right)$	μ	σ^2

4.3. 协方差

4.3.1. 协方差的概念

设(X,Y)是二维随机变量，如果$E[(X-EX)(Y-EY)]$存在，则称它为随机变量X与Y的协方差，记作$Cov(X,Y)=E[(X-EX)(Y-EY)]$.

4.3.2. 协方差的计算公式

对任意两个随机变量X,Y，有

(1) $Cov(X,Y)=E(XY)-E(X)E(Y)$；

(2) $D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$.

4.3.3. 协方差的性质

(1) $Cov(X,Y)=Cov(Y,X)$；

(2) $Cov(X,X)=D(X)$；

(3) $Cov(X,c)=0$ （c为任意常数）

(4) $Cov(aX,bY)=ab\,Cov(X,Y))$ （a,b是常数）；

(5) $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$；

4.4. 相关系数

4.4.1. 相关系数的定义

对于随机变量X与Y，如果X,Y的方差都不等于零，则称$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$为随机变量$X$与$Y$的相关系数，记作$\rho_{XY}$（简记为$\rho$）

4.4.2. 相关系数的性质

(1) $|\rho_{XY}|\leqslant 1$；

(2) 当$\rho_{XY}=0$时，称X与Y不相关.
若X,Y相互独立，则$\rho_{XY}=0$，反之不成立；
(3) $|\rho|=1 \Leftrightarrow X$与$Y$以概率1线性相关.
若∃常数$a>0$，b，使$P\{X=aY+b\}=1$，则$\rho_{XY}=1$，此时称X，Y正相关；
若∃常数$a<0$，b，使$P\{X=aY+b\}=1$，则$\rho_{XY}=-1$，此时称X，Y负相关；

4.4.3. 随机变量X,Y等价的五个结论

(1) X与Y不相关
(2) $\rho_{XY}=0$
(3) $Cov(X,Y)=0$
(4) $E(XY)=E(X)E(Y)$
(5) $D(X\pm Y)=D(X)+D(Y)$

5. 大数定律与中心极限定理

5.1. 切比雪夫不等式

设随机变量X具有$E(X)$和$D(X)$，则任给ε > 0，有

$$P\{|X-E(X)|\geqslant\varepsilon\}\leqslant\frac{D(X)}{\varepsilon^2}，或P\{|X-E(X)|<\varepsilon\}\geqslant 1-\frac{D(X)}{\varepsilon^2}.$$

5.2. 大数定律

5.2.1. 依概率收敛

设a是一个常数，X_n为一随机变量序列，$\forall\varepsilon>0$，$\exists\, P\{|X_n-a|<\varepsilon\}=1$或$P\{|X_n-a|\geqslant\varepsilon\}=0$，则称$\{X_n\}$依概率收敛于$a$，记为$X_n\xrightarrow{P}a$.

5.2.2. 伯努利大数定律（即频率依概率收敛于概率）

设n_A是n次独立重复试验中事件A发生的次数，$P(A)=p$，则$\forall\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{n_A}{n}-p\right|<\varepsilon\right\}=1或\lim_{n\to\infty}P\left\{\left|\frac{n_A}{n}-p\right|\geqslant\varepsilon\right\}=0$$

5.2.3. 辛钦大数定律

设随机变量$X_1, X_2, \cdots, X_n$相互独立，服从同一分布（任意分布），且具有相同的数学期望$E(X_k) = \mu$，$\overline{X} = \frac{1}{x}\sum_{k=1}^{n} X_k$，则$\forall\ \varepsilon > 0$，有

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

5.2.4. 切比雪夫大数定律

设随机变量$X_1, X_2, \cdots, X_n$相互独立，数学期望$E(X_k)$，方差$D(X_k)$均存在，且方差有公共的上界即$\forall k$，$D(X_k) \leqslant C$，$\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$，则$\forall\ \varepsilon > 0$，有

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

其中$\mu = \frac{1}{n}\sum_{k=1}^{n} E(X_k)$.

5.3. 中心极限定理

5.3.1. 独立同分布中心极限定理

设$X_1, X_2, \cdots, X_n, \cdots$ 是独立同分布的随机变量序列，$EX_i = \mu, DX_i = \sigma^2, i = 1, 2, \cdots$，则随机变量$Y_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$的分布函数$F_n(x), \forall x \in R$，有

$$\lim_{n\to\infty} F_n(x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t$$

即，当n充分大时，$\sum_{i=1}^{n} X_i$近似地服从以它的均值为均值，它的方差为方差的正态分布，即正态分布为$N\left(n\mu, n\sigma^2\right)$.

5.3.2. 拉普拉斯中心极限定理

设n_A表示n重伯努利试验中时间A出现次数，$P(A)=p$，则随机变量$Y_n=\dfrac{n_A-np}{\sqrt{np(1-p)}}$的分布函数$F_n(x)$，$\forall x\in R$，有

$$\lim_{n\to+\infty}F_n(x)=\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t$$

即，当n充分大时，n_A近似地服从它的均值为均值，它的方差为方差的正态分布，即正态分布$N(np,np(1-p))$.

6. 数理统计的基本概念

6.1. 常用统计量

设X是总体，其均值$E(X)=\mu$，方差$D(X)=\sigma^2$.

6.1.1. 样本均值

$\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$，样本均值的期望与方差分别为$E(\overline{X})=\mu$，$D(\overline{X})=\dfrac{\sigma^2}{n}$.

6.1.2. 样本方差

$S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2=\dfrac{1}{n-1}\left[\sum\limits_{i=1}^{n}X_i^2-n\overline{X}^2\right]$，样本方差的期望与方差分别为$E\left(S^2\right)=\sigma^2$，$D\left(S^2\right)=\dfrac{2\sigma^4}{n-1}$

样本标准差：$S=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}$

6.2. 抽样分布

6.2.1. 三大分布

(1) 卡方分布：$\chi^2\sim\chi^2(n)$

- (a) 定义：设$X_1,X_2,\cdots,X_n$是相互独立同服从$N(0,1)$分布，则称$\chi^2=X_1^2+X_2^2+\cdots+X_n^2$的分布为服从自由度为$n$的卡方分布，记为$\chi^2\sim\chi^2(n)$.
- (b) α上分位点：如果$P\left(\chi^2>\chi_\alpha^2\right)=\alpha(0<\alpha<1)$，则称点$\chi_\alpha^2(n)$为$\chi^2-\chi^2(n)$上的$\alpha$上分位点.在$\chi^2-\chi^2(n)$上，$\chi_a^2(n)$与$\chi_{1-\alpha}^2(n)$之间没有关系.
- (c) 性质

(i) 设$\chi_1^2 - \chi^2(n_1)$，$\chi_2^2 - \chi^2(n_2)$，且χ_1^2, χ_2^2相互独立，则$\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$.该性质可推广到任意有限个的情况.

(ii) 设$\chi^2 \sim \chi^2(n)$，则$E\left(\chi^2\right) = n$，$D\left(\chi^2\right) = 2n$.

(2) t分布：$T \sim t(n)$

(a) 定义：设$X \sim N(0,1)$，$Y \sim \chi^2(n)$，且X, Y相互独立，则称$T = \dfrac{X}{\sqrt{\dfrac{Y}{n}}}$的分布为服从参数为$n$的$t$分布，记为$T \sim t(n)$.

(b) t分布的概率密度函数为偶函数.

(c) α上分位点：如果$P(T > t_n(n)) = \alpha(0 < \alpha < 1)$，则称点$t_\alpha(n)$为$T \sim t(n)$上的$\alpha$上分位点.

(d) 关系：$t_{1-\alpha}(n) = -t_\alpha(n)$

(3) F分布：$F \sim F(n_1, n_2)$

(a) 定义：设$X \sim \chi^2(n_1)$，$Y \sim \chi^2(n_2)$，且X, Y相互独立，则称$F = \dfrac{\frac{X}{n_1}}{\frac{Y}{n_2}}$的分布为服从参数为$(n_1, n_2)$的$F$分布，记为$F \sim F(n_1, n_2)$

(b) α分位点：如果$P(F > F_\alpha(n_1, n_2)) = \alpha(0 < \alpha < 1)$，称点$F_\alpha(n_1, n_2)$为$F \sim F(n_1, n_2)$上的$\alpha$分位点.

(c) 性质：设$F \sim F(n_1, n_2)$，则称$\dfrac{1}{F} \sim F(n_1, n_2)$

(d) 关系：$F_{1-\alpha}(n_1, n_2) = \dfrac{1}{F_\alpha(n_2, n_1)}$

6.3. 关于样本的分布

1. 设单正态总体$X \sim N\left(\mu, \sigma^2\right)$，其样本为$X_1, X_2, \cdots, X_n$，样本均值为$\overline{X}$，样本方差为$S^2$.

a. $\dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$ （$\overline{X} - N\left(\mu, \dfrac{\sigma^2}{n}\right)$）

b. $\dfrac{\sum\limits_{i=1}^{n} (X_i - \mu)}{\sigma^2} \sim \chi^2(n)$

c. $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ （$\dfrac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{\sigma^2} \sim \chi^2(n-1)$），且$\overline{X}$与$S^2$相互独立

d. $\dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$

2. 设双正态总体$X \sim N\left(\mu_1, \sigma_1^2\right)$，其样本为$X_1, X_2, \cdots, X_{n_1}$，样本均值为$\overline{X}$，样本方差为$S_1^2$，$Y - N\left(\mu_2, \sigma_2^2\right)$，其样本为$Y_1, Y_2, \cdots, Y_{n_2}$，样本均值为$\overline{Y}$，样本方差为$S_2^2$，且样本$X_1, X_2, \cdots, X_{n_1}$，$Y_1, Y_2, \cdots, Y_{n_2}$相互独立.

a. $\dfrac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} \sim F(n_1-1, n_2-1)$

b. $\dfrac{n_2\sigma_2^2 \sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2}{n_1\sigma_1^2 \sum\limits_{i=1}^{n_2}(X_i-\mu_2)^2} \sim F(n_1, n_2)$

c. $\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$

d. 当$\sigma_1^2=\sigma_2^2$未知时，$\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2)$其中$S_w=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

7. 参数估计与假设检验

7.1. 点估计

设$X_1, X_2, \cdots, X_n$为来自总体X的样本，若将样本的某个函数$\hat{\theta}(X_1, X_2, \cdots, X_n)$作为总体分布中未知参数$\theta$的估计，则称$\hat{\theta}$为$\theta$的点估计量. 在抽样后，$\hat{\theta}$的值$\hat{\theta}(x_1, x_2, \cdots, x_n)$称为$\theta$的估计值.

点估计分为矩估计量与最大似然估计法.

7.1.1. 矩估计法

设总体X的分布中含有m个未知参数$\theta_1, \theta_2, \cdots, \theta_m$，令

$$E\left(X^k\right) = \frac{1}{n}\sum_{i=1}^{n} X_i^k, k = 1, 2, \cdots, m$$

则由上述方程所求得的解：$\hat{\theta}_k(X_1, X_2, \cdots, X_n), k = 1, 2, \cdots, m$ 称为未知参数θ_K的矩估计量，简称矩估计，只要掌握$m = 1, 2$的情形.

7.1.2. 最大似然估计法

设总体X的概率密度为$f(x, \theta_1, \cdots, \theta_m)$（若$X$为离散型，则用分布律代替），$\theta_1, \cdots, \theta_m$为未知参数，记$x_1, \cdots, x_n$为样本$X_1, \cdots, X_n$的观测值，则称

$$L(x_1, \cdots, x_n; \theta_1, \cdots, \theta_m) = \prod_{i=1}^{n} f(x_i; \theta_1, \cdots, \theta_m)$$

为似然函数，若有$\hat{\theta}(x_1, \cdots, x_n)(i = 1, \cdots, m)$使得

$$L(x_1, \cdots, x_n; \hat{\theta}_1, \cdots, \hat{\theta}_m) = \max_{(\theta_1, \cdots, \theta_m)} L(x_1, \cdots, x_n; \theta_1, \cdots, \theta_m)$$

则称$\hat{\theta}_i(x_1, \cdots, x_n)$为$\theta_1$的最大似然估计值，而将$\hat{\theta}_i(X_1, \cdots, X_n)$称为$\theta_i$最大似然估计量（$i = 1, 2, \cdots, m$）.只要掌握$m = 1, 2$的情形.

对于最大似然估计，其求解步骤为：

(1) 写出似然函数$L(x_1, x_2, \cdots, x_n; \theta_1, \cdots, \theta_m) = \prod_{i=1}^{n} f(x_i; \theta_1, \cdots, \theta_m)$；

(2) 取对数$\ln L$；

(3) 求偏导数$\dfrac{\partial \ln L}{\partial \theta_i}, i = 1, 2, \cdots, m$;

(4) 判断方程（组）$\dfrac{\partial \ln L}{\partial \theta_i} = 0$是否有解，若有解，则其解即为所求最大似然估计，若无解，则最大似然估计常在θ_i的边界上达到。

7.2. 估计量的评选标准

7.2.1. 无偏性

设$\hat{\theta}_i(X_1, \cdots, X_n)$为$\theta$的估计量，若$E(\hat{\theta}) = \theta$，则称$\hat{\theta}$为$\theta$的无偏估计.

7.2.2. 有效性

设$\hat{\theta}_1, \hat{\theta}_2$均为$\theta$的无偏估计，若$D(\hat{\theta}_1) < D(\hat{\theta}_2)$，则称$\hat{\theta}_1$比$\hat{\theta}_2$有效.

7.2.3. 一致性

设$\hat{\theta}$为θ的估计量，若对任意的$\varepsilon > 0$，有

$$\lim_{n\to\infty} P(|\hat{\theta}-\theta| < \varepsilon) = 1$$

即$\hat{\theta} \xrightarrow{P} \theta$, $n \to \infty$，则称$\hat{\theta}$为θ的一致估计量（或相合估计量）.

7.3. 区间估计

设θ为总X总的分布中的位置参数，$X_1, X_2, \cdots, X_n$为取自X的样本，若存在两个统计两$\hat{\theta}_1(X_1, X_2, \cdots, X_n)$，$\hat{\theta}_2(X_1, X_2, \cdots, X_n)$使得对给定的$a(0 < a < 1)$，有

$$P\{\hat{\theta}_1 \leqslant \theta \leqslant \hat{\theta}_2\} = 1 - a$$

则称$[\hat{\theta}_1, \hat{\theta}_2]$为$\theta$的置信度为$1-a$的置信区间，$\hat{\theta}_1, \hat{\theta}_2$分别称为置信下限和置信上限.

正态总体未知参数的置信区间如下页表.

待估参数		抽样分布	双侧置信区间
μ	σ^2已知	$U = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$	$(\overline{X} - \mu_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \overline{X} + \mu_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}})$ $P\{\lvert U\rvert \geqslant \mu_{\frac{\alpha}{2}}\} = \alpha$
	σ^2未知	$T = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$(\overline{X} - t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}, \overline{X} + t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}})$ $P\{\lvert T\rvert \geqslant t_{\alpha}\} = \alpha$

σ^2	μ已知	$W' = \frac{1}{\sigma^2}\sum_{j=1}^{n}(X_i - u)^2 \sim \chi^2(n)$	$P\left\{W' \geq \chi^2_{\frac{a}{2}}(n)\right\} = P\left\{W' \leq \chi^2_{1-\frac{a}{2}}(n)\right\} = \frac{a}{2}$
	μ未知	$W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\frac{(n-1)S^2}{\chi^2_{\frac{a}{2}}}(n-1), \frac{(n-1)S^2}{\chi^2_{1-\frac{a}{2}}(n-1)}\right)$
$\mu - \mu_2$	σ_1^2, σ_2^2 已知	$U = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N$	$\left((\overline{X}_1 - \overline{X}_2) - u_{\frac{0}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},\right.$ $\left.(\overline{X}_1 - \overline{X}_2) + u_{\frac{a}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},\right)$ $P\{\lvert U\rvert \geq u_{\frac{a}{2}}\} = a$
	已知 $\sigma_1^2 = \sigma_2^2$ $= \sigma^2$ 但σ^2 未知	$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} - t(n_1 + n_2 - 2$ $S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$	$\left((\overline{X}_1 - \overline{X}_2) - t_{\frac{a}{2}} \cdot (n_1 + n_2 - 2) \cdot S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right.$ $\left.(\overline{X}_1 - \overline{X}_2) + t_{\frac{a}{2}} \cdot (n_1 + n_2 - 2) \cdot S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}},\right)$ $P\{\lvert T\rvert \geq t_a\} = a$
$\frac{\sigma_1^2}{\sigma_2^2}$		$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} - F(n_1 - 1, n_2 - 1)$	$(\frac{1}{F_{\frac{a}{2}}(n_1 - 1, n_2 - 1)} \cdot \frac{S_1^2}{\sigma_2^2},$ $F_{\frac{a}{2}}(n_2 - 1, n_1 - 1) \cdot \frac{S_1^2}{S_2^2})$ $P\{F \geq F_{\frac{a}{2}}(m_1 - 1, n_2 - 1)\} = \frac{a}{2},$ $P\left\{\frac{1}{F} \geq F_{\frac{a}{2}}(n_2 - 1, n_1 - 1)\right\} = \frac{a}{2},$

7.4. 假设检验

7.4.1. 显著性检验的基本思想

为了对总体的分布类型或分布中的未知参数做出推断，首先对它们提出一个假设H_0，然后再H_0为真的条件下，通过选取恰当的统计量来构造一个小概率事件，若在一次试验中，小概率事件居然发生了，就完全有理由拒绝H_0的正确性，否则没有充分理由拒绝H_0的正确性，从而接受H_0，这就是显著性检验的基本思想.

7.4.2. 假设检验的基本步骤

(1) 由实际问题提出原假设H_0（与备选假设H_1）；
(2) 选取适当的统计量，并在H_0为真的条件下确定该统计量的分布；
(3) 根据问题要求确定显著性水平α（一般题目中会给定），从而得到拒绝域；
(4) 由样本观测值计算统计量的观测值，看是否属于拒绝域，从而对H_0做出判断.

7.4.3. 两类错误

当H_0本来是正确的，但检验后作出了拒绝H_0的判断，这种错误称为第一类错误，也称拒真错误；当H_0本来是不正确的，但检验后作出了接受H_0的判断，这种错误称为第二类错误，也称受伪错误。

7.4.4. 正态总体未知参数的假设检验（检验水平α）

	原假设H_0	H_0下的检验统计量及分布	H_0的拒绝域

一个正态分布	$\mu=\mu_0$ (σ^2已知)	$U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$	$\lvert u\rvert=\left\lvert\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right\rvert\geqslant u_{\frac{\alpha}{2}}$
	$\mu=\mu_0$ (σ^2未知)	$T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$	$\lvert t\rvert=\left\lvert\dfrac{\overline{x}-\mu_0}{S/\sqrt{n}}\right\rvert\geqslant t_{\frac{\alpha}{2}}(n-1)$
	$\sigma^2=\sigma_0^2$ (μ已知)	$W=\sum_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma_0}\right)^2\sim\chi^2(n)$	$w=\sum_{i=1}^{n}\left(\dfrac{x_i-\mu}{\sigma_0}\right)^2\geqslant\chi^2_{\frac{\alpha}{2}}(n)$ 或$w\leqslant\chi^2_{1-\frac{\alpha}{2}}(n)$
	$\sigma^2=\sigma_0^2$ (μ未知)	$W=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$	$w=\dfrac{(n-1)S^2}{\sigma_0^2}\geqslant\chi^2_{\frac{\alpha}{2}}(n-1)$ 或$w\leqslant\chi^2_{1-\frac{\alpha}{2}}(n-1)$
两个正态分布	$\mu_1-\mu_2=\delta$ (σ_1^2,σ_2^2已知)	$U=\dfrac{\overline{X_1}-\overline{X_2}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$	$\lvert u\rvert=\left\lvert\dfrac{\overline{X_1}-\overline{X_2}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\right\rvert\geqslant u_{\frac{\alpha}{2}}$

	$\mu_1-\mu_2=\delta$ (σ_1^2,σ_2^2未知，但 $\sigma_1^2=\sigma_2^2$)	$T=\dfrac{\overline{X_1}-\overline{X_2}-\delta}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$ $S_W^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-1}$	$\lvert t\rvert=\left\lvert\dfrac{\overline{X_1}-\overline{X_2}-\delta}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\right\rvert\geqslant t_{\frac{\alpha}{2}}(n_1+n_2-2)$
	$\sigma_1^2=\sigma_2^2$ (μ_1,μ_2未知)	$F=\dfrac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1)$	$f=\dfrac{S_1^2}{S_2^2}\geqslant F_{\frac{\alpha}{2}}(n_1-1,n_2-1)$ 或$f\leqslant F_{\frac{\alpha}{2}}^{-1}(n_2-1,n_1-1)$