# 概率统计

邱一航
520030910155

# 随机事件与概率

随机现象：无法预言，多种结果，大量试验下有统计规律性

随机实验：相同条件下可重复，无法预知，多种结果

随机事件：（随机试验中）可能发生也可能不发生的事件。（A, B, C）

将随机事件用集合描述以便研究

样本空间 $\Omega = \bigcup_i A_i$（随机试验的所有直接结果合体组成的集合）

$\omega \in \Omega$

元素：样本点（基本事件）

任意随机事件 $\longrightarrow$ 样本空间的子集（$A \subset \Omega$）

e.g. {死了，没死}

$\{0, 1, \dots, +\infty\}$ 个人在教室里.

$\{0, 1, \dots, Max\}$ 个人在教室里.

$\{t \in (0, +\infty)\}$.

特殊情况：必然事件（$\Omega$），不可能事件（$\varnothing$）.

- 关系与运算

维恩图即可表示.

$A \subset B$ $\rightarrow$ A的发生必定导致B的发生

$A \subset B, B \subset A \Rightarrow A = B$

$A \cup B$（A和B的和事件）$\longrightarrow$ A和B两个事件至少有一个会发生的事件

$\bigcup_{i=1}^{n} A_i$ $\rightarrow$ $A_1, \dots A_n$ n个事件.（n可以为 $\infty$）

$A \cap B$ 或 $AB$（A和B的积事件）$\longrightarrow$ A和B两个事件同时会发生的事件

$A_1, \dots A_n$，及 $n = \infty$ 同理.

$A - B$（A和B的差事件）$\longrightarrow$ A发生同时B不发生的事件.

A与B $\boxed{\text{互不相容（互斥）}}$ $\longleftarrow$ $A \cap B = \varnothing$.

A和B不可能同时发生[redacted].

[redacted] $A_1, \dots A_n$ 两两互不相容 $\longleftarrow$ $A_iA_j = \varnothing$（$\forall \{i, j\} \subset \{1, 2, \dots n\}$）

A与B $\boxed{\text{相互对立（互逆）}}$ $\longleftarrow$ $A \cap B = \varnothing$，$A \cup B = \Omega$

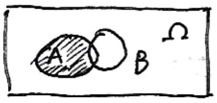
A和B不可能同时（不）发生[redacted].

记 $B = \overline{A}$，称B为A的对立事件（逆事件）

"A不发生" "$\overline{A}$发生"

运算规律： 1) $\bar{\bar{A}} = A$ （重余律）

2) $A\bar{B} = A - B = A - AB$




3). 交换律、结合律、分配律

4) De'Morgan 定律： $\overline{A\cup B} = \bar{A}\bar{B}$ $\qquad \overline{AB} = \bar{A}\cup\bar{B}$

（反演律）

A和B中至少一个发生的逆事件 ↓ 至多0个发生 ↓ 都不发生

A和B都发生的逆事件（至少2个） ↓ 至多1个发生 ↓ 至少1个不发生

e.g. A：通过α B：通过β. C：通过γ

基本事件： $2^3 = 8$ （ $ABC$, $\underbrace{\bar{A}BC,\ A\bar{B}C,\ AB\bar{C}}_{3}$, $\underbrace{\bar{A}\bar{B}C,\ A\bar{B}\bar{C},\ \bar{A}B\bar{C}}_{3}$, $\bar{A}\bar{B}\bar{C}$ ）

不多于2项通过 $\overline{ABC} = \bar{A}\cup\bar{B}\cup\bar{C}$

- 古典概型

古典概型 { 基本事件总数有限；各基本事件等可能发生 ↓ 要求："均匀"

$$P(A) = \frac{M\ \text{(A中包含的基本事件个数)}}{N\ \text{(基本事件总数)}}$$

n个不可分辨球，m个可分辨盒. (1) 最多1个球/盒. 指定的k个盒里都有球的概率

(2) 不限制盒中球数. 指定的k个盒里无球的概率. （设每球的放法均可能）

[解] (1) $\dfrac{C_{m-k}^{n-k}}{C_m^n}$

(2) 解 ⟶ 隔板法 $\quad 0\cdots0 \mid 0\cdots0 \mid\mid\mid 0\cdots0 \mid \cdots \mid$

↑ 分隔盒子 共 $n+m-1$ 个符号

只要确定0就确定一种放置方法 ⟶ 样本空间大小 $|\Omega| = C_{n+m-1}^{n}$

不妨令指定k个盒为前k个盒无球. 即前(k)个位置全是"|". 剩余位置 $n+m-k-1$

$|A_2| = C_{n+m-k-1}^{n}$

$$\therefore P(A_2) = \frac{C_{n+m-k-1}^{n}}{C_{n+m-1}^{n}}. \qquad \square$$

注意：球不可分辨！因此$m^n$←是有问题的.

改：$n$个球完全随机地投入$m$个盒．求指定$k$个盒中无球的概率．

[解] ~~此[illegible]~~．

唯一可能的方式

此时基本事件是某个球落在某个盒中．　$|\Omega| = m^n$　　$|A_3| = (m-k)^n$

每个球$m$种选择

$$P(A_3) = \frac{(m-k)^n}{m^n}$$

□

易出问题 → 基本事件的设计

- 几何概型

几何概型 $\begin{cases} \text{样本空间对应于有限空间 } \Omega. & \sim \text{“有限”} \\ \text{样本点落在 } \Omega \text{ 内任何区域 } G \text{ 的概率与 } G \text{ 的测度成正比，与其所处位置无关} & \searrow \text{“等可能性”} \end{cases}$

$(l, S, V, \cdots)$

$$P(A) = \frac{A\text{所对应区域的测度}}{\Omega\text{的测度}}$$

e.g. 甲乙两船彼此无通讯，24h内停靠某码头．若两船分别需停靠1h、2h．求一昼夜内任一船到达时需停靠．

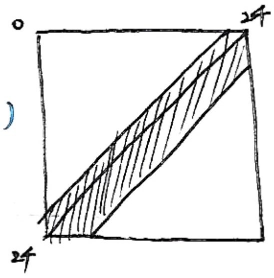
[解] 甲到达：$x$．$0 \le x < 24$　；　乙到达：$y$．$0 \le y < 24$．

$\Omega = \{(x,y) \mid 0 \le x < 24,\ 0 \le y < 24\}$

等待现象：$y-1 \le x < y+2$

$\left( \begin{cases} y = x. \\ y > x: & y \le x+1 \\ y < x: & y \ge x-2 \end{cases} \right)$

等号有无没影响（几何概型．）




$$P(A) = 1 - P(\bar{A}) = 1 - \frac{S(\bar{A})}{S(\Omega)} = 1 - \frac{\frac{1}{2}\cdot(23^2+22^2)}{24^2}$$

□

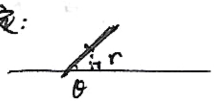
e.g. Buffon 投针.　　针长 $l$. 平行线无数条. 等距 $a$.　$l<a$.

[解] $l<a \Rightarrow$ 针一定在某二条（相邻）平行线范围内.　可以改变这二条平行线的选取，使离针的中心最近的平行线恰为三条平行线中居中的那条

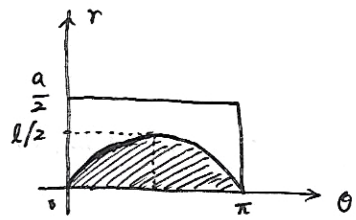
选这三条


如何定位针？ $\dashrightarrow (r,\theta)$

相交的判定:




$\frac{l}{2}\sin\theta \geq r$




样本空间　$\Omega = \left\{(r,\theta) \,\middle|\, 0\leq r\leq \frac{a}{2},\ 0\leq\theta\leq\pi\right\}$

$$P(A) = \frac{S(A)}{S(\Omega)} = \frac{\frac{l}{2}\cdot\int_0^{\pi}\sin x\,dx}{\frac{a}{2}\cdot\pi} = \frac{2l}{a\pi}$$

随机试验
用于计算 $\pi$.（在含参变量积分搞出 $\pi$ 的算法之前）

* 贝特朗悖论

- 频率与概率的统计定义

频率.　定义略

$\boxed{\text{概率的统计定义}}$　随机事件在重复的 $n$ 次随机试验中出现 $m$ 次. 若随 $n$ 增大，$P(A)=\frac{m}{n}$ 始终稳定在某个数值 $p$ $(0\leq p\leq 1)$ 附近，称 $p$ 为概率（统计定义）.

- 概率的公理化定义

发现三种模型下，概率都满足:

$0\leq P(A)\leq 1$

$P(\Omega)=1$

$A_1, A_2\cdots$ 两两互不相容. $P\left(\bigcup_{i=1}^{+\infty}A_i\right)=\sum_{i=1}^{+\infty}P(A_i)$

$\boxed{\text{概率的公理化定义}}$

设 E 是一个随机试验，$\Omega$ 为其对应的样本空间. 若存在映射 $\Omega\to\mathbb{R}$，（法则使对 E 中任一事件 A 赋予一个实数）. 记为 $P(A)$. 满足:

(1) $\forall A\subseteq\Omega$. $P(A)\geq 0$　(2) $P(\Omega)=1$　(3) 对 $A_1, A_2, \cdots\cdots$ 两两互不相容事件.（无限个）

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

则称 $P(A)$ 为事件A的 $\boxed{\text{概率}}$.

推导得出性质：

1) $P(\phi)=0$

[证] $\infty$个两两互斥事件 $A_1, A_2, \cdots$ 使 $\bigcup^{\infty} A_i=\Omega$ 则 $\overset{P(\Omega)}{1}=\sum_{i=1}^{\infty}P(A_i)$

另$\infty$个：$\phi, A_1, A_2, \cdots$ $1=P(\phi)+\sum_{i=1}^{\infty}P(A_i)$

$\therefore P(\phi)=0$ □

2) 有限可加性：$n$个 $(n<+\infty)$ 两两互不相容事件

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

[证] $\infty$个两两互斥事件 $B_1, B_2, \cdots$

另$\infty$个 $A_1, \cdots A_n, B_1, B_2, \cdots$

相减可得.

3) $P(\bar{A})=1-P(A)$ [证] (2) 可证.

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

[证] $P(A)=P[(A-AB)\cup AB]=P(A-AB)+P(AB)$
$=P(A-B)+P(AB)$. □

5) $A\subseteq B \Rightarrow P(A)\leqslant P(B)$

6) $P(A\cup B)=P(A)+P(B)-P(AB)$ 易证、拆成两两互斥事件

7) 容斥原理：$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{l=1}^{n}\sum_{1\leq k_1<k_2<\cdots<k_l\leq n}P\left(\bigcup_{i=1}^{l}A_{k_i}\right)\cdot(-1)^{l+1}.$$

e.g. 1, 2, … 9 中可重复地取出 $n$个数 $(n\geqslant 2)$. 求$n$次所取数的乘积为10倍数的概率.

[解] 原事件记为 A. $A_1$：$n$次取到数字中有偶数. $A_2$：$n$次取到数字中有5.

则 $A=A_1A_2$. $\bar{A}=\overline{A_1A_2}=\bar{A_1}\cup\bar{A_2}$

$$P(\bar{A})=P(\bar{A_1}\cup\bar{A_2})=P(\bar{A_1})+P(\bar{A_2})-P(\bar{A_1}\cap\bar{A_2})=\frac{5^n}{9^n}+\frac{8^n}{9^n}-\frac{4^n}{9^n}$$

$$\therefore P(A)=1-P(\bar{A})=1-\left(\frac{5}{9}\right)^n-\left(\frac{8}{9}\right)^n+\left(\frac{4}{9}\right)^n.$$

e.g. 配对问题. $A_1, \cdots A_n$. $a_1, \cdots a_n$ 随机配对. 至少出现一对 $(A_i, a_i)$ 的概率.

[解] 原事件记为B. 记 $B_i$：$A_i$, $a_i$ 配对的概率. 则有 $B=\bigcup_{i=1}^{n}B_i$

$$\text{则 } P(B)=P\left(\bigcup_{i=1}^{n}B_i\right)=\sum_{i=1}^{n}P(B_i)-\sum_{1\leq i<j\leq n}P(B_iB_j)+\sum_{1\leq i<j<k\leq n}P(B_iB_jB_k)-\cdots+(-1)^{n+1}P(B_1\cdots B_n)$$

↓

变为抽签问题

$$=\sum_{i=1}^{n}\frac{1}{n}-C_n^2\cdot\frac{1}{n(n-1)}+C_n^3\cdot\frac{1}{n(n-1)(n-2)}\cdots+(-1)^{n+1}\cdot\frac{1}{n!}$$

$$=1-\frac{\cancel{n(n-1)}}{2!}\cdot\frac{1}{\cancel{n(n-1)}}+\frac{\cancel{n(n-1)(n-2)}}{3!}\cdot\frac{1}{\cancel{n(n-1)(n-2)}}-\cdots+(-1)^{n+1}\cdot\frac{1}{n!}$$

$$=\sum_{i=1}^{n}(-1)^{i+1}\cdot\frac{1}{i!}$$

e.g. 证明 $C_N^n=\sum_{k=s_1}^{s_2}C_M^kC_{N-M}^{n-k}$ 其中 $s_1=\max\{0, n-N+M\}$. $s_2=\min\{n, M\}$

[证] 即证 $$1=\frac{\sum_{k=s_1}^{s_2}C_M^kC_{N-M}^{n-k}}{C_N^n}$$ 即"$N$个球. 不放回地拿$n$个. $N$个球中有$M$个红球, 其余白球. 其中$k$个是红球的概率为 $\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}$"

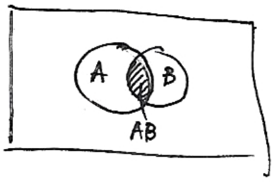
(拿出的球)

另一方面. $k \le M$. $k \le n$ $\longrightarrow$ $k \le \min\{M, n\}$

$n-k \le N-M$. $\longrightarrow$ $k \ge \max\{0,\ n-N+M\}$ □

- 条件概率

记 A事件发生的条件下B发生的概率为 $P(B|A)$ . 称为$\boxed{\text{条件概率}}$




$$P(B|A) := \frac{P(AB)}{P(A)} \qquad (P(A)>0) \quad \text{易理解.}$$

[证明] ~~$P(AB) \le P(A) \Rightarrow P(B|A)$~~ 下面证明 $P(\cdot|A)$ 是概率.

$$P(AB) \ge 0 \Rightarrow P(B|A) \ge 0.$$

$$P(A|A) = \frac{P(A)}{P(A)} = 1.$$

一系列两两互斥事件 $B_1, B_2, \cdots B_\infty$. $P\left(\bigcup_{i=1}^{\infty} B_i \,\middle|\, A\right) = \frac{P\left(\bigcup_{i=1}^{\infty} B_i\right)}{P(A)} = \frac{\sum_{i=1}^{\infty} P(B_i)}{P(A)} = \sum_{i=1}^{\infty} P(B_i|A)$

$\therefore$ $P(\cdot|A)$ 拥有概率的一切性质 □

注: $P(B|A) + P(B|\bar{A}) \ne 1$. $P(A|B) + P(\bar{A}|B) = 1$.

$\boxed{\text{~~Bayes~~ 乘法公式}}$ $\underbrace{P(AB) = P(A)P(B|A)}_{\text{乘法公式}} = P(B)P(A|B) \quad (P(A)>0,\ P(B)>0)$

n个随机事件乘法公式:

$$\begin{aligned} P(A_1A_2\cdots A_n) &= P(A_1A_2\cdots A_{n-1})\,P(A_n|A_1A_2\cdots A_{n-1}) \\ &= P(A_1\cdots A_{n-2})\,P(A_{n-1}|A_1\cdots A_{n-2})\,P(A_n|A_1\cdots A_{n-1}) \\ &= \underbrace{P(A_1)}_{A_1}\,\underbrace{P(A_2|A_1)}_{A_1\text{条件下}A_2}\cdots\cdots \underbrace{P(A_n|A_1\cdots A_{n-1})}_{A_1\cdots A_{n-1}\text{条件下}A_n} \end{aligned}$$

$A_1, A_2$条件下$A_3$

e.g. 13个人5元. 13个人10元. 每人收5元. 不会遇到找不出零钱的概率

[解] 记第i个人来时不会找不出钱的事件为 $A_i$ (拿着10元) $A = A_1A_2\cdots A_{13}$

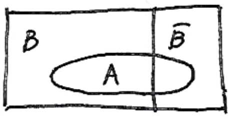
$$P(A) = P(A_1)\,P(A_2|A_1)\,P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})$$

$$= \frac{13}{14}\cdot\frac{12}{13}\cdot\frac{11}{12}\cdot\cdots\cdot\frac{1}{2} = \frac{1}{14}$$ □

$\{0,1,\cdots,13\}$ $\{1,\cdots,13\}-1$
$=\{0,\cdots,12\}$
只剩这些5元

• 全概率与 Bayes 公式

两种观察：

① 

② $\bar{B} \to A$，$\bar{B} \to A$

$\boxed{\text{全概率公式}}$ 记 $B_1, \cdots B_n$ 为样本空间 $\Omega$ 的一个划分，即 $\bigcup_{i=1}^{n} B_i = \Omega$ 且 $B_iB_j = \emptyset$ $(\forall \{i,j\} \subseteq \{1,\cdots,n\})$

则对任意事件A 有 $P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$

对任意事件A，$P(A)>0$.

$\boxed{\text{Bayes 公式}}$ $$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n} P(B_j)P(A|B_j)} \quad (i=1,2,\cdots,n)$$

结果 $\xrightarrow{\text{回看}}$ 条件

实际上only要 $A \subset \bigcup_{i=1}^{n} B_i$ 即可

不一定要是划分.

先验概率 . 后验概率 → 得到A发生信息后进行修正和重新估计 $P(B_i|A)$

↓

在A是否发生无信息的条件下对致其发生原因 $B_i$ 概率的事先估计

$P(B_i)$

e.g. m人传球. 第1次甲传出给其他(m-1)个人. 求第n次由甲传出的概率.

[解] $P_i$：第i次甲传出. $P_i = P_{i-1} \cdot 0 + \frac{1}{m-1}(1-P_{i-1}) = \frac{1}{m-1}(-P_{i-1}+1)$ …… □

• 事件独立性

$\boxed{\text{相互独立}}$ 1) $P(A|B) = P(A|\bar{B}) = P(B)$ 则A, B相互独立

2) A的发生不影响B，反之亦然.

即 $P(A)>0$ 时 $P(B|A) = P(B)$

$P(B)>0$ 时. $P(A|B) = P(A)$ $\xrightleftharpoons[]{\text{需要}P(A)>0\text{且}P(B)>0}$ $P(AB) = P(A)P(B)$

(未考虑 $\Omega$ 和 $\emptyset$)

} 性质

定义：设A, B为两事件. 若 $P(AB) = P(A)P(B)$，则称 A, B.

性质3) 四对事件 $(A,B)$. $(A,\bar{B})$. $(\bar{A},B)$, $(\bar{A},\bar{B})$. 任一对相互独立，则其它三对也相互独立

互不相容事件（非 $\emptyset$ 和 $\Omega$）一定不能相互独立. 易证.

不可能~ 必然~

[推广] 三元：两两相互独立且三者彼此独立 $\begin{cases} P(AB) = P(A)P(B) \\ P(AC) = P(A)P(C) \\ P(BC) = P(B)P(C) \end{cases}$

$P(ABC) = P(A)P(B)P(C)$

n元：……

相互独立的 $n$ 个事件分成几组（各组之间无交叉部分），各组内随意运算所得事件彼此也相互独立。

e.g. $A_1, A_2, \cdots A_6 \Rightarrow A_1 \cap A_2 \cup \overline{A_3}$、$A_4 A_5$、$\overline{A_6}$ 相互也独立。

# 随机变量及其分布

- 随机变量

~~E为一样本空间~~，

E为一随机事件，$\Omega$ 为其对应的样本空间，若 $\forall \omega \in \Omega$ 按一定法则存在 $X(\omega) \in \mathbb{R}$ 与之对应，则称X为$\boxed{\text{随机变量}}$。 $X: \Omega \to \mathbb{R}$ （X本身取值随机、不确定）

$\begin{cases} \text{离散型} \\ \text{非离散型} \begin{cases} \text{连续型随机变量} \\ \cdots \end{cases} \end{cases}$

符号：$X, Y, Z$ 或 $\xi, \eta, \zeta$

- 分布函数（Cumulative Distribution Function）

设X为随机变量，对每个实数 $x$，随机事件 $\{X \leqslant x\}$ 的概率 $P(X \leqslant x) =: F(x)$，称 $F(x)$。

* Notice: $X: \Omega \to \mathbb{R}$ $\quad F: \mathbb{R} \to [0,1]$

则有 $P(a < X \leqslant b) = P(X \leqslant b) - P(X \leqslant a) = F(b) - F(a)$

$P(X > a) = 1 - P(X \leqslant a) = 1 - F(a)$

$P(X = x_0) = \lim\limits_{\Delta x \to 0^+} P(x_0 - \Delta x < X \leqslant x_0) = \lim\limits_{\Delta x \to 0^+} F(x_0) - F(x_0 - \Delta x) = F(x_0) - F(x_0^-)$

左极限

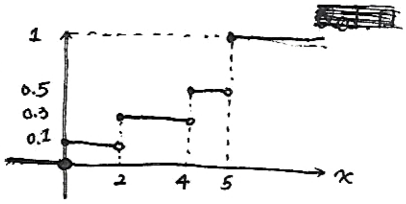
性质：(1) 单调不减；(2) $\lim\limits_{x \to +\infty} F(x) = 1$，$\lim\limits_{x \to -\infty} F(x) = 0$ 记作 $F(+\infty) = 1$，$F(-\infty) = 0$。

(3) 右连续，即 $\lim\limits_{x \to x_0^+} F(x) = F(x_0)$

[证明] $F(x_0 + \Delta x) = P(X \leqslant x_0 + \Delta x) = F(x_0) + P(x_0 < X \leqslant x_0 + \Delta x) = F(x_0) + \int_{x_0}^{x_0+\Delta x} P(X = u)\,du$

$P(X = u) \in [0,1]$，$\therefore \Delta x \to 0^+$ 时，$\int_{x_0}^{x_0+\Delta x} P(X = u)\,du \to 0$。$\therefore F(x_0 + \Delta x) \to F(x_0)$ $(\Delta x \to 0^+)$ □

e.g. $Z = \begin{cases} 5 & 50\% \\ 4 & 20\% \\ 2 & 20\% \\ 0 & 10\% \end{cases}$ $\qquad F(x) = \begin{cases} 0 & (x < 0) \\ 0.1 & (0 \leqslant x < 2) \\ 0.3 & (2 \leqslant x < 4) \\ 0.5 & (4 \leqslant x < 5) \\ 1 & (x \geqslant 5) \end{cases}$




- 离散型随机变量

X的取值有限个或可列无穷个，称X $\boxed{\text{离散型随机变量}}$

常用描述：$\boxed{\text{分布律}}$（概率分布） $P(X = x_k) =: p_k$ $\quad \begin{cases} \text{非负} \\ \sum \text{全} = 1. \end{cases}$ * 可列个 → 有界正项级数

$P(X = x_k) = F(x_k) - F(x_k^-) = F(x_k) - F(x_{k-1})$ （定义 $x_0 = -\infty$）

* 左极限也常写作 $F(x-0)$。

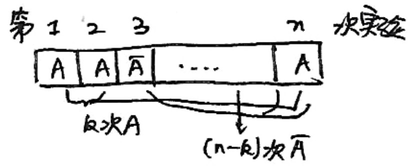
0-1分布 $\begin{array}{c|c} 0 & 1 \\ \hline p & 1-p \end{array}$ 有时分布律写作 $p^k(1-p)^k \quad (k\in\{0,1\})$

二项分布 $\longrightarrow$ 将一些事件视作相互独立，简化模型（n重 Bernoulli 试验模型）

n重 Bernoulli 概型：相同独立的重复n次，每次都只对事件A进行考察。 $P(A)=p,\ 0<p<1$.

整个概型考察 A在n次试验中发生次数的概率。

$X=0,1,\cdots,n$




$$\boxed{P(X=k)=p^k(1-p)^{n-k}\cdot C_n^k}$$ （二项分布）

记作 $X\sim B(n,p)$

Binomial
二项分布

e.g. 从甲和乙两组各抓一个人对抗，甲队人员获胜概率0.6，以下三种方案（均以胜出人数多的一组为胜）何种对乙最有利？

(1) 各取三人； (2) 各取五人； (3) 各取七人

乙获胜人数为X.

[解] (1) $P(乙)=P(X=2)+P(X=3)=0.6^1\cdot(1-0.6)^2\cdot C_3^2+(1-0.6)^3\cdot C_3^3=0.352$

(2) $P(乙)=P(X=3)+P(X=4)+P(X=5)=C_5^3\cdot(1-0.6)^3\cdot0.6^2+C_5^4(1-0.6)^4\cdot0.6^1+C_5^5(1-0.6)^5=0.317$

(3) $P(乙)=P(X=4)+P(X=5)+P(X=6)+P(X=7)=\cdots=0.290$

[另解] 比赛次数.

(2) $P(乙)=(0.4)^3+\underset{\uparrow}{C_3^2}(0.4)^3\times0.6+\underset{\uparrow}{C_4^2}(0.4)^3\times0.6^2=0.317$

前三次有2次乙胜 $\longleftarrow$ 最终都是乙胜

e.g. 从1~10中有放回地抽取数字(5个)，将取出的数字从小到大排列，正中间是4的概率.

[解] 记排序后的数字为 $X_1,X_2,X_3,X_4,X_5$. 所求即 $\boxed{P(X_3=4)=P(X_3\leq4)-P(X_3\leq3)}$

$(X_1\leq X_2\leq X_3\leq X_4\leq X_5)$

只要取出的数字有至少3个小于等于4即可

X：抽取的5个数字中 $\{\leq4\}$ 发生的次数. $X\sim B(5,\frac{4}{10})$

$P(X_3\leq4)=P(X=3)+P(X=4)+P(X=5)=C_5^3(\frac{4}{10})^3(\frac{6}{10})^2+C_5^4(\frac{4}{10})^4(\frac{6}{10})+C_5^5(\frac{4}{10})^5$

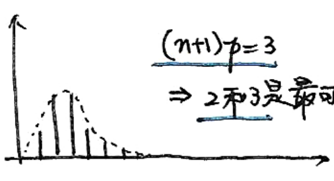
$\xi$: ···· $\leq3$ ···· $\xi\sim B(5,\frac{3}{10})$

$P(X_3\leq3)=P(\xi=3)+P(\xi=4)+P(\xi=5)=C_5^3$ ····

$P(X_3=4)=$ ——— . $\square$

二项分布取值情况

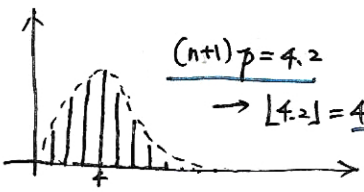
$B(8,\frac{1}{3})$



$(n+1)p=3$ (2个)

$\Rightarrow$ 2和3是最可能取值（最概然）

$B(20,0.2)$



$(n+1)p=4.2$ (1个)

$\rightarrow \lfloor4.2\rfloor=4\rightarrow$ 唯一最概然取值

证明：考察 $\frac{P(X=k+1)}{P(X=k)}$ !

$\boxed{\text{Poisson 分布}}$　　9:00—12:00 到来的客户数X的概率分布？

（泊松分布）

1) 分成很多（N个）区间. 每个区间来0个人/1个人. 否则再分. → $N^+$个区间

2) 各个区间内客户数相互独立

3) 每个区间出现1个客户的概率与区间长度成正比. 不妨设为 $\frac{\lambda}{N}$. λ为常数 ← 这三个小时的平均客户数

$$X_N \sim B(N, \frac{\lambda}{N}) \qquad P(X_N=k) = C_n^k p^k (1-p)^{n-k}$$

$$\lim_{N\to\infty} P(X_N=k) = \lim_{N\to\infty} C_N^k \left(\frac{\lambda}{N}\right)^k \left(1-\frac{\lambda}{N}\right)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}$$

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

因此 $\left\{\frac{\lambda^k}{k!}e^{-\lambda}\right\}$ 可作为分布律. $(k=0,1,2,\cdots)$

因此，称 $\boxed{P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda} =: P(\lambda)}$ 为 Poisson 分布. 记作 $X\sim P(\lambda)$

$\boxed{\text{Poisson 定理}}$　设 $\lim_{n\to\infty} np_n = \lambda$. 则对固定的k有 $\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$

（二项分布推到极限成为 Poisson分布）　　$(k=0,1,2,\cdots)$

[证明] 记 $np_n =: \lambda_n$.

$$\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = \lim_{n\to\infty}\frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k} = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\cdot\frac{\lambda_n^k}{k!}\cdot\left[\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}}\right]^{-\lambda_n}\left(1-\frac{\lambda_n}{n}\right)^{-k}$$

$$= \lim_{n\to\infty} 1\cdot\frac{\lambda^k}{k!}\cdot e^{-\lambda_n}\cdot 1 = \frac{\lambda^k}{k!}e^{-\lambda}$$

应用：若 $X\sim B(n,p)$, n较大. p较小，$np=\lambda$ 适中时. 认为近似Poisson分布.

即 $C_n^k p^k(1-p)^{n-k} \approx e^{-\lambda}\frac{\lambda^k}{k!}$　　书中附录 λ: [redacted] 0.1～10.

e.g. 宇宙飞船. 宇宙粒子进舱概率分布：$\underset{\text{进舱粒子数}}{X} \sim P(\lambda)$. $\overset{\text{X个}}{\text{粒子}}$落到重要部位的概率为p. 求落到重要部位的粒子数概率分布.

[解] 记落到……粒子数为Y.　　$Y=m \mid X=k \sim B(k,p)$

$$\text{全概率公式}\quad P(Y=m) = \sum_{m=0}^{+\infty} P(X=k)\,P(Y=m\mid X=k)$$

$$= \sum_{k=m}^{+\infty} e^{-\lambda}\frac{\lambda^k}{k!}C_k^m p^m(1-p)^{k-m}$$

$$= \sum_{k=m}^{+\infty}\frac{\lambda^k}{\cancel{k!}}\frac{\cancel{k!}\,p^m}{m!(k-m)!}e^{-\lambda}(1-p)^{k-m} = \frac{e^{-\lambda}\lambda^m p^m}{m!}\sum_{k=m}^{+\infty}\frac{[\lambda(1-p)]^{k-m}}{(k-m)!} = \frac{\lambda^m p^m}{m!}e^{\lambda(1-p)}e^{-\lambda}$$

$$= \frac{\lambda^m p^m}{m!}e^{-p\lambda} = \frac{(p\lambda)^m}{m!}e^{-p\lambda}$$　　也是 Poisson 分布

即 $Y \sim P(p\lambda)$

* $\boxed{\text{巴斯卡分布}}$（负二项分布）

重复射击目标，直到命中 r 次所需的射击次数 X.

$P(X=k) = C_{k-1}^{r-1} p^{r-1}(1-p)^{k-r} \cdot p = C_{k-1}^{r-1} p^r (1-p)^{k-r}$. 特别地，$r=1$ 称$\boxed{\text{几何分布}}$ $\underline{(1-p)^{k-1} \cdot p}$

$\boxed{\text{超几何分布}}$ 与 $\boxed{\text{二项分布}}$ 的关系：$a+b \to \infty$. $\frac{a}{a+b} \to p$. 此时"放不放回"影响不大.

（有放回地拿球）（无放回地拿球）

$$\frac{C_a^k C_b^{n-k}}{C_{a+b}^n} \to C_n^k p^k (1-p)^{n-k}.$$

（超几何分布）

- 连续型随机变量

频率直方图：高度是$\underline{\text{频率/区间长度}}$！$\longrightarrow$ 为了让面积是频率.

概率密度函数.

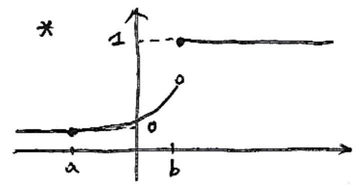
$\boxed{\text{连续型随机变量}}$ X：随机变量. $F(x)$：分布函数. 若 $\exists f(x)$ s.t. $F(x) = \int_{-\infty}^{x} f(u)\,du$ $(-\infty < u < +\infty)$

则称 X.

概率密度

处处连续、几乎处处可导

[redacted] *



随机变量取值连续. $[a,b]$

$\underline{\text{但不是连续型随机变量！}}$（找不到概率密度函数）

注意：$f(x)$ 不唯一. e.g. $f_1(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$ 和 $f_2(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$

$\begin{cases} f(x) \ge 0 \\ \int_{-\infty}^{+\infty} f(x)\,dx = 1. \end{cases}$ $\longrightarrow$ 可作为概率密度.

概率密度 $\ne$ 概率. 概率 $= \int_{\text{区间}}$ 概率密度·微元！（面积！） $f(x_0)\Delta x \doteq P(x_0 < X \le x_0 + \Delta x)$ $(\Delta x \to 0)$

$$P(X=a) = \lim_{\Delta x \to 0} \int_{a-\Delta x}^{a} f(u)\,du = 0.$$

* 不可能事件 $\overset{\times}{\rightleftarrows}$ 概率为 0

必然事件 $\overset{\times}{\rightleftarrows}$ 概率为 1.

e.g. $f(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$.

- $x = \frac{1}{2}$：不是不可能事件
- $x \ne \frac{1}{2}$：不是必然事件.
- $x = 2$：不可能事件
- [redacted] $x \in \mathbb{R}$：必然事件.

$\boxed{\text{均匀分布}}$：概率密度是常值. Uniform Distribution

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \text{（可取可不取）} \\ 0. & \text{otherwise} \end{cases}$$ 记作 $X \sim U(a,b)$.

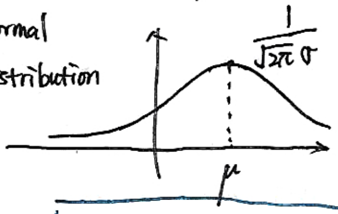
$\boxed{\text{指数分布}}$：Exponential Distribution.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0. & \text{otherwise} \end{cases} \Rightarrow F(x) = \begin{cases} 0. & (x < 0) \\ 1 - e^{-\lambda x} & (x \ge 0) \end{cases}$$

与 Poisson 分布的关系

$\boxed{\text{正态分布}}$ $N(\mu, \sigma^2)$

Normal Distribution




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

$P(X \le x) = P(X \ge 2\mu - x)$ ← 对称性

~~$P(|X| \le x)$~~ 标准正态分布 $P_n(y)$. → 分布函数 $\Phi(y)$ $y = \frac{x-\mu}{\sigma}$

$P_n(|X| \le x) = P_n(-x \le X \le x) = 2P(X \le x) - 1.$

$= 2\Phi(x)$

特别地, $X^* \sim N(0,1)$ 分布函数记作 $\Phi(x)$.

std Normal Distribution

* 若 $f(\mu+x) = f(\mu-x)$

其对称性质类似 $N(\mu, \sigma^2)$

* $\int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)}{2\sigma^2}} dx$ 没有对应的初等函数 ~ 借助 $\Phi(x)$ 查表.

$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ $F(\mu + \sigma x) = \Phi(x)$

* $3\sigma$ 原理: $|X-\mu| < 3\sigma$

$\Rightarrow P_N(-3 < X^* < 3)$

$= 0.9974$

- 从已知概率分布推得相关事件概率分布.

离散型: 简单 对应的加一加完事. → 分布列

(或积分)

连续型: 需要概率密度描述.

$F_Y(y) = P(Y \le y) = P(f(X) \le y) = P(X \,\square\, x)$ 可求 ← 分布函数.

* 概率密度: 保留原函数的积分形式. 或直接对分布函数求导

e.g. $Y = aX + b$. $X \sim N(\mu, \sigma^2) \Rightarrow Y \sim N(a\mu + b, \ \sigma^2|a|^2)$

我们要求正态分布中的"$\sigma$" $>0$.

$Y = f(X)$. $f$ 单调 e.g. $Y = 1 - e^{-\lambda X}$. $X \sim E(\lambda)$. $Y \sim$ 均匀分布 $(0,1)$

分布函数. $= U(0,1)$

$\boxed{\text{定理}}$ 若 $Y = F_X(x)$. ⟶ $Y \sim U(0,1)$.

→ Monte-Carlo 仿真算法

$\boxed{\text{定理}}$ 已知 $X: f_X(x)$. $g(x)$ 为 R 上严格单调可导函数. $Y = g(X)$

则 $Y$ 的概率密度为

$$f_Y(y) = \begin{cases} |h'(y)| f_X(h(y)) & \alpha < y < \beta \\ 0 & \text{otherwise} \end{cases}$$

其中 $h(y) = g^{-1}(x)$. $\{\alpha, \beta\} = \{g(-\infty), g(+\infty)\}$.

证明: ……

# 多维随机变量及其分布

- 二维随机变量及其分布函数

$\boxed{\text{二维随机变量}}$ $\Omega$：随机试验样本空间 $\forall \omega \in \Omega$. 唯一的 $\exists (X(\omega), Y(\omega))$ 与之对应 $\in \mathbb{R}^2$.

称 $(X, Y)$ 向量. 故也称二维随机向量.

$\boxed{\text{分布函数}}$ $F_{X,Y}(x,y) := P(X \leq x, Y \leq y) := P[(X \leq x) \cap (Y \leq y)]$

这是个记号！

称二维随机变量 $(X, Y)$ 的联合分布函数.

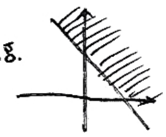
* $0 \leq F(x,y) \leq 1$, $F(+\infty, +\infty) = 1$, $F(-\infty, -\infty) = 0$. $F(x, -\infty) = 0$. ($x$ 可取 $+\infty$！)

$F(-\infty, y) = 0$ ($y$ 可取 $+\infty$！)

* 对 $x, y$ 都单调不减（需固定一个，对单变量单调不减）

* 对 $x\ y$ 单变量右连续（…………右连续）

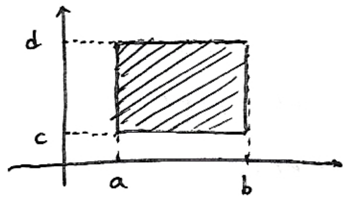
e.g.



$f(x,y) = \begin{cases} 1, & x+y \geq 1 \\ 0, & x+y < 1 \end{cases}$

不满足第四条性质！

* $\forall a < b,\ c < d$.



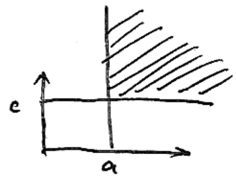
$P(a < X \leq b,\ c < X \leq d)$

$= F(b, d) - F(a, d) - F(b, c) + F(a, c) \geq 0$.

满足以上四条性质的函数能作为二维随机变量的联合分布函数.

* $P(X > a, \cancel{X} > c) = 1 + F(a, c) - F(a, +\infty) - F(+\infty, c)$

$\neq 1 - F(a, c)$！



$\boxed{\text{边缘分布}}$ $F_X(x) = P(X \leq x) = F(x, +\infty) := \lim\limits_{y \to +\infty} F(x, y)$ 称 $X$ 的边缘分布函数

e.g. $F(x,y) = A\left(B + \arctan\frac{x}{2}\right)\left(C + \arctan\frac{y}{2}\right)$ 是联合分布函数

(1) $A, B, C$. (2) $F_X$, $F_Y$. (3) $P(X > 2)$.

[解]

(1)

$F(+\infty, +\infty) = A\left(B + \frac{\pi}{2}\right)\left(C + \frac{\pi}{2}\right) = 1$.

$F(-\infty, -\infty) = A\left(B - \frac{\pi}{2}\right)\left(C - \frac{\pi}{2}\right) = 0$

$F(+\infty, -\infty) = A\left(B + \frac{\pi}{2}\right)\left(C - \frac{\pi}{2}\right) = 0$

$F(-\infty, +\infty) = A\left(B - \frac{\pi}{2}\right)\left(C + \frac{\pi}{2}\right) = 0$

$\Rightarrow \begin{cases} A = \frac{1}{\pi^2} \\ B = \pi/2 \\ C = \pi/2 \end{cases}$

* 实际上，题已告知是联合分布函数. 列三个就行，不用一一验证性质.

(2) $F_X(x) = F(x, +\infty)$

$= \frac{1}{\pi}\left(\frac{\pi}{2} + \arctan\frac{x}{2}\right)$

$F_Y(y) = F(+\infty, y)$

$= \frac{1}{\pi}\left(\frac{\pi}{2} + \arctan\frac{y}{2}\right)$

(3) $P(X > 2) = 1 - F_X(2) = 1 - \frac{3}{4} = \frac{1}{4}$.

□

$\boxed{\text{二维离散型随机变量}}$ $(X,Y) \in \{(x_i, y_j),\ i,j=1,2,\dots\}$

$P(X=x_i, Y=y_j) = p_{ij}$ 分布列/分布律 ~ 联合分布表

性质：1) $p_{ij} \geq 0\ \forall i \forall j$　　2) $\sum_i \sum_j p_{ij} = 1$.

| $p_{ij}$ | | X: $x_1$ | $\cdots$ | $x_i$ | $\cdots$ | $p_{\cdot j} = \sum_i p_{ij}$ |
|---|---|---|---|---|---|---|
| Y | $y_1$ | $p_{11}$ | $\cdots$ | $p_{i1}$ | $\cdots$ | $p_Y(1)$ |
| | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ |
| | $y_j$ | $p_{1j}$ | $\cdots$ | $p_{ij}$ | $\cdots$ | $p_Y(j)$ |
| | $\vdots$ | $\vdots$ | | $\vdots$ | $\ddots$ | $\vdots$ |
| $p_{i\cdot} = \sum_j p_{ij}$ | | $p_X(1)$ | $\cdots$ | $p_X(i)$ | $\cdots$ | $\sum_i \sum_j p_{ij} = 1$ |

边缘分布

边缘分布

由联合分布函数确定联合分布列：

$p_{ij} = F(x_i, y_j) - F(x_i - 0, y_j) - F(x_i, y_j - 0) + F(x_i - 0, y_j - 0)$

不好求.

e.g. 仨球，仨盒. X：落入1盒球数　Y：落入2盒球数. 等可能落入. 求(1) 联合分布律, 边缘分布律.
(2) $P(X=Y)$, $P(Y>X)$. (3) $F(x,y)$

[解] (1)

| Y \ X | 0 | 1 | 2 | 3 | |
|---|---|---|---|---|---|
| 0 | 1/27 | $C_3^1/27$ | $C_3^2/27$ | $C_3^3/27$ | |
| 1 | 3/27 | 6/27 | 3/27 | 0 | $\vdots$ |
| 2 | 3/27 | 3/27 | 0 | 0 | |
| 3 | 1/27 | 0 | 0 | 0 | |
| | | $\cdots$ | | | 1 |

(未化简)

① 古典概型算

② $P(X=i, Y=j) = \underbrace{P(X=i)}_{\text{二项分布}} P(Y=j \mid X=i)$

$= C_3^i \left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^{3-i} \cdot C_{3-i}^j \left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^{3-i-j}$

$= C_3^i \cdot \frac{\cancel{2^{3-i}}}{3^3} \cdot C_{3-i}^j \cdot \frac{1}{\cancel{2^{3-i}}}$

$= \frac{C_3^i C_{3-i}^j}{3^3}$

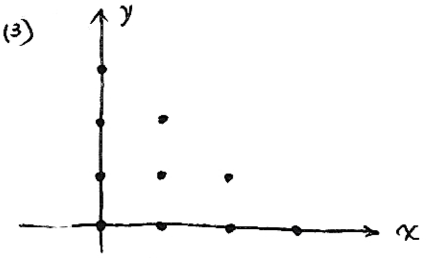
X, Y 分布一致
⇓
称 $\boxed{\text{X, Y 同分布}}$
v.s. X, Y 相等 (X=x 则 Y=y)

边缘分布你加一加就好. 略.

(2) $P(X=Y) = \frac{1}{27} + \frac{6}{27} = \frac{7}{27}$.

$P(Y>X) = \frac{3}{27} + \frac{3}{27} + \frac{3}{27} + \frac{1}{27} = \frac{10}{27}$

(3)




⇒ …… 略.

□

* 边缘分布一样时，联合分布依然有多种不同可能

$\boxed{\text{二维连续型随机变量}}$ 设二维随机变量 X, Y …… $F(x,y)$

存在非负可积函数 $f(x,y)$ s.t. $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv$

称 $\boxed{\text{联合概率密度}}$

性质：$f(x,y) \geq 0,\ \forall x \forall y$　　$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(u,v)\,du\,dv = 1$.

$P(\text{点})=0$.　$P(\text{直线})=0$.　$P(\text{可求长曲线})=0$.

边缘概率分布：$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\int_{-\infty}^{+\infty}f(u,v)\,du\,dv=\int_{-\infty}^{x}du\underbrace{\boxed{\int_{-\infty}^{+\infty}f(u,v)\,dv}}_{=f_X(u)}$

$$f_X(x)=\frac{dF_X(x)}{dx}=\int_{-\infty}^{+\infty}f(x,y)\,dy$$

y同理，略.

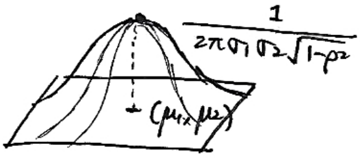
**均匀分布**　面积 A.　$f(x,y)=\begin{cases}\frac{1}{A}. & (x,y)\in G\\ 0 & \text{otherwise}\end{cases}$ ：　$X,Y\sim U(G)$

* 当G是边平行坐标轴的矩形时，其边缘分布依然均匀分布.

* G无限制时，未必！　e.g.  边缘分布不均匀.

**二维正态分布**　$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}$

vs 一维正态：$\frac{1}{\sqrt{2\pi}\sigma}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$




~~n维正态分布的矩阵形式：~~ $\cancel{f(x)=\frac{1}{(\sqrt{2\pi})^n|\Sigma|}\exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]}$

~~对二维~~ $\cancel{\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}}$

($\Sigma$ 是n阶正定方阵)

记作 $(X,Y)\sim U(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$.　$|\rho|\leq 1$

n维正态分布的矩阵形式：$f(x)=\frac{1}{(\sqrt{2\pi})^n|\Sigma|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]$

对二维的情况即：$\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$　$x=\begin{pmatrix}x\\ y\end{pmatrix}$　$\mu=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$

e.g. $\int_{-\infty}^{+\infty}e^{-x^2}dx$ ：① $=\sqrt{\iint e^{-(x^2+y^2)}dx\,dy}=\sqrt{\pi}$.

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

边缘分布：$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,dy=\int_{-\infty}^{+\infty}\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\left(\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2+\left(\frac{x-\mu_1}{\sigma_1}\right)^2-\rho^2\frac{(x-\mu_1)^2}{\sigma_1^2}\right)\right]$

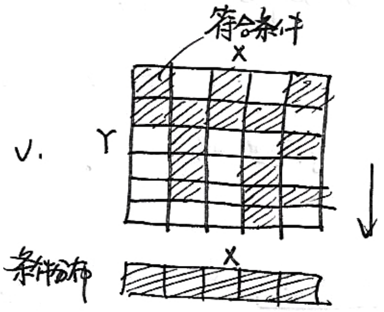
$$=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\underbrace{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)\sigma_2^2}\left(y-\left(\mu_2+\rho\frac{\sigma_2(x-\mu_1)}{\sigma_1}\right)\right)^2}}_{=1}$$

$$=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$　也是正态分布.

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

e.g. $f(x,y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}(1+\sin x\sin y)$ $\longrightarrow$ 边缘分布也正态

- 随机变量的条件分布




离散型二维随机变量的条件分布： 对应条件下的值在联合分布中加一加 $\longrightarrow$ v.

$P(Y=y_j\mid X=x_i)\geq 0$. $\sum_{j=1}^{n}P(Y=y_j\mid X=x_i)=1$.

连续型二维随机变量的条件分布

记号： $f_{Y|X}(y|x)$： $X=x$条件下Y的条件概率密度

自变量 ↑ 固定，但x只有在有分布处才有意义

$$f_{Y|X}(y|x)=\frac{f(x,y)}{\int_{-\infty}^{+\infty}f(x,y)\,dy}=\frac{f(x,y)}{f_X(x)}$$ $>0$ 才有意义 $\Rightarrow$ 类似乘法公式： $f(x,y)=f_{Y|X}(y|x)f_X(x)$

连续点（因为边界点原本就可在可不在 $\longrightarrow$ 概率密度不唯一）

正态分布的条件分布

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

[redacted]

- 随机变量的独立性

$P(X\leq x, Y\leq y)=P(X\leq x)P(Y\leq y)$ 即 $F(x,y)=F_X(x)F_Y(y)$.

$\Leftrightarrow$ $P(\{X\in D_1\}\{Y\in D_2\})=P(X\in D_1)P(X\in D_2)$.

二维正态分布 $(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$ $\underline{\rho=0}$ ~ 相互独立

- 多维随机变量的函数的分布

离散型： 显然，列表!

$X\sim B(n_1,p)$. $Y\sim B(n_2,p)$. X.Y相互独立 则 $X+Y\sim B(n_1+n_2,p)$

$X\sim P(\lambda_1)$. $Y\sim P(\lambda_2)$ X.Y相互独立 $X+Y\sim P(\lambda_1+\lambda_2)$

连续型： 概率密度 [redacted]. $f(x,y)$

e.g. $Z=X+Y$.

$$F_Z(z)=P(X+Y\leq z)=\iint_{x+y\leq z}f(x,y)\,dx\,dy=\int_{-\infty}^{+\infty}dx\int_{-\infty}^{z-x}f(x,y)\,dy$$

$$=\int_{-\infty}^{+\infty}dx\int_{-\infty}^{z}f(x,v-x)\,dv=\int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty}f(x,v-x)\,dx\right]dv$$

$$\therefore f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\,dx$$

类似可得：$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)\,dx = \int_{-\infty}^{+\infty} f(z-y,y)\,dy$

$(Z=X+Y)$

若X, Y相互独立，则 $f_Z(z)=\int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,dx =: f_X * f_Y(z)$

卷积！

$$aX+bY+c=Z \Rightarrow f_Z(z)=\frac{1}{|b|}\int_{-\infty}^{+\infty} f\left(t, \frac{z-at-c}{b}\right) dt$$

$$=\frac{1}{|a|}\int_{-\infty}^{+\infty} f\left(\frac{z-bt-c}{a}, t\right) dt$$

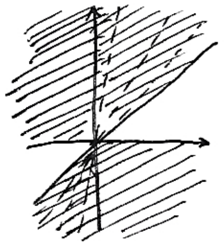
$X_1, \cdots X_n$ 相互独立. $(\rho=0)$ $X_i \sim N(\mu_i, \sigma_i^2) \Rightarrow \sum X_i \sim N(\sum\mu_i, \sum\sigma_i^2)$

线性函数 $\sim N($线$\mu_i$, 线$\sigma_i^2)$

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

商的分布：$f(x,y)$ $Z=X/Y$.

$$F_Z(z)=P(Z\le z)=P\left(\frac{X}{Y}\le z\right)=\int_{-\infty}^{+\infty}dy\int_{\square}^{\square} f(x,y)\,dy=\int_0^{+\infty}dy\int_{-\infty}^{zy} f(x,y)\,dx + \int_{-\infty}^{0}dy\int_{zy}^{+\infty} f(x,y)\,dx$$



$dx=d(yz)=y\,dz$

$$=\int_0^{+\infty}dy\int_{-\infty}^{z} f(zy,y)\,dz\cdot y + \int_{-\infty}^{0}dy\int_{z}^{-\infty} f(zy,y)\cdot y\,dz$$

$$=\int_{-\infty}^{z}dz\int_0^{+\infty} f(zy,y)\,y\,dy + \int_{-\infty}^{z}dz\int_{-\infty}^{0} f(zy,y)\,(-y)\,dy$$

$$=\int_{-\infty}^{z}dz\ \underbrace{\int_{-\infty}^{+\infty} f(zy,y)\,|y|\,dy}_{\boxed{f_Z(z)=\int_{-\infty}^{+\infty} f(zy,y)\,|y|\,dy}}$$

$Z'=Y/X$. $\boxed{f_{Z'}(z)=\int_{-\infty}^{+\infty} f(x,zx)\,|x|\,dx}$

$Z=X^2+Y^2$ $\boxed{f_Z(z)=\begin{cases}0 & z<0\\ \frac{1}{2}\int_0^{2\pi} f(\sqrt{z}\cos\theta,\sqrt{z}\sin\theta)\,d\theta. & z\ge 0\end{cases}}$ （极坐标）

若 $X_i\sim N(0,1)$. 则 $Z=\sum X_i^2$ 服从的分布为 $f_Z(z)=\begin{cases}0 & ,z<0\\ \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} z^{\frac{n}{2}-1} e^{-\frac{z}{2}} & ,z\ge 0\end{cases}$

$\downarrow$

自由度为n的 $\chi^2$ 分布

$$\Gamma(x)=\int_0^{+\infty} e^{-t}t^{x-1}\,dt$$

极值分布：$F_{Max}(u) = P(Max \leq u) = F(u,u)$

$F_{Min}(v) = 1 - P(X>v, Y>v) = \cdots\cdots$

若 $X_1, \cdots X_n$ 相互独立. $X_i \sim F_i(x)$, $i = 1, 2, \cdots, n$

$M = \max\{X_1, \cdots X_n\}$ $N = \min\{X_1, \cdots X_n\}$.

$$F_M(z) = P(X_1 \leq z, \cdots X_n \leq z)$$
$$= F_1(z) F_2(z) \cdots\cdots F_n(z)$$

$$F_N(z) = 1 - P(X_1 > z, \cdots X_n > z) = 1 - (1 - F_1(z))(1 - F_2(z)) \cdots\cdots (1 - F_n(z))$$

即：$F_M(z) = \prod F_i(z)$ $F_N(z) = 1 - \prod (1 - F_i(z))$

# 随机变量的数字特征

- 数学期望

离散型随机变量：$P(X=x_i)=p_i$ ．$i=1,2,\cdots$ 若 $\sum_{i=1}^{+\infty} x_i p_i$ 绝对收敛．

称其和为 X 的数学期望．记作 $\mathbb{E}X = E(X) = \sum_{i=1}^{+\infty} x_i p_i$ ．

（考虑到有限情况：$\mathbb{E}X = \sum_{i\in I} x_i p_i$ ）

连续型随机变量：密度函数为 $f(x)$．若广义积分 $\int_{-\infty}^{+\infty} x f(x)\,dx$ 绝对收敛

则称其为 X 的数学期望．记作 $\mathbb{E}X = E(X) = \int_{-\infty}^{+\infty} x f(x)\,dx$ ．

e.g. Cauchy 分布：$f(x)=\dfrac{1}{\pi(1+x^2)}$ ．其广义积分 $\int_{-\infty}^{+\infty} |x| f(x)\,dx = 2\int_0^{+\infty} \dfrac{x}{\pi(1+x^2)}\,dx$ 发散

数学期望不存在．（ $\int_{-N}^{N^2}$ ．和 $\int_{-N^2}^{N}$ $N\to+\infty$ ．）

随机变量函数的数学期望．

$g(x)$ 函数．离散型：若 $\sum_{i\in I} g(x_i)p_i$ 绝对收敛．$=: \mathbb{E}g(X)$

随机型：若 $\int_{-\infty}^{+\infty} g(x)f(x)\,dx$ 绝对收敛．$=: \mathbb{E}g(X)$．

二维随机变量．$\mathbb{E}g(X,Y) = \sum_{i\in I}\sum_{j\in J} g(x_i,y_j)\,p_{ij}$ （若存在）

$\mathbb{E}g(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,f(x,y)\,dx\,dy$ （若存在）

$\mathbb{E}X^k$ k阶原点矩

$\mathbb{E}|X|^k$ k阶绝对原点矩

$\mathbb{E}(X-\mathbb{E}X)^k$ k阶中心矩 $(\mu=\mathbb{E}X)$

特别地，$\mathbb{D}X = \mathbb{E}(X-\mathbb{E}X)^2$ 称方差．

$\mathbb{E}(X^k+Y^l)$ k+l阶混合原点矩

$\mathbb{E}((X-\mathbb{E}X)^k(Y-\mathbb{E}Y)^l)$ k+l阶混合中心矩

$k=l=1$ 时：k+l阶混合原点矩称二阶…

k+l ……中心矩称二阶….

即 协方差．

$\mathbb{E}\dfrac{(X-\mathbb{E}X)(Y-\mathbb{E}Y)}{\sqrt{\mathbb{D}X}\sqrt{\mathbb{D}Y}}$ ：相关系数．

数学期望的性质：

1) $\mathbb{E}\,Const = Const$

2) $\mathbb{E}aX = a\mathbb{E}X$

3) $\mathbb{E}(\mu X+\eta Y) = \mu\mathbb{E}X+\eta\mathbb{E}Y$

↑ 3条 线性性

4) X, Y 相互独立 $\Rightarrow$ $\mathbb{E}(XY)=(\mathbb{E}X)(\mathbb{E}Y)$

5) $P(X\geq a)=1 \Rightarrow \mathbb{E}X\geq a$

6) $P(X\leq b)=1 \Rightarrow \mathbb{E}X\leq b$．

3) 证明：以连续型为例。只证 $\mathbb{E}X+\mathbb{E}Y=\mathbb{E}(X+Y)$.

$$\mathbb{E}X+\mathbb{E}Y=\iint x f(x,y)\,dx\,dy+\iint y f(x,y)\,dx\,dy=\iint (x+y) f(x,y)\,dx\,dy=\mathbb{E}(X+Y)$$

4) 证明：连续型。 $\mathbb{E}(XY)=\iint xy\, f(x,y)\,dx\,dy=\iint xy\, f_X(x) f_Y(y)\,dx\,dy=\int_{-\infty}^{+\infty} f_X(x)\,x\,dx\int_{-\infty}^{+\infty} f_Y(y)\,y\,dy$

$=(\mathbb{E}X)(\mathbb{E}Y)$

e.g. $\mathbb{E}X\,\mathbb{E}Y=\mathbb{E}(XY)$ 但不独立：$(X,Y)\sim U(D)$， ✓ $D=\{(x,y)\mid x^2+y^2\le 1\}$.

e.g. $n$ 个球落入 $n$ 个 $V=\infty$ 的盒子，求空着盒子数 $X$ 的数学期望.

$X=X_1+X_2+\cdots+X_n$

| $X_i=0$ | $1$ |
|---|---|
| $1-\left(\frac{n-1}{n}\right)^n$ | $\left(1-\frac{1}{n}\right)^n$ |

$$\mathbb{E}X=\sum \mathbb{E}X_i=\sum\left(1-\frac{1}{n}\right)^n=n\left(1-\frac{1}{n}\right)^n$$
$$=\frac{(n-1)^n}{n^{n-1}}.$$

- 方差

若 $(\mathbb{E}(X-\mathbb{E}X)^2)$ 存在，$=: \mathbb{D}X=D(X)$，称方差，也记作 $Var(X)$

"距离"，因绝对值不好处理.

Deviation　　Variance

称 $\sqrt{\mathbb{D}X}$ 为标准差.　　衡量取值偏离的情况.

离散型 $\mathbb{D}X=\sum_{i\in I}(x_i-\mathbb{E}X)^2 p_i$　　连续型：$\mathbb{D}X=\int_{-\infty}^{+\infty}(x-\mathbb{E}X)^2 f(x)\,dx$

单位$^2$　单位

$$\boxed{\mathbb{D}X=\mathbb{E}(X^2)-(\mathbb{E}X)^2}$$

证明：$\mathbb{D}X=\mathbb{E}((X-\mathbb{E}X)^2)=\mathbb{E}(X^2-2X\cdot\mathbb{E}X+(\mathbb{E}X)^2)=\mathbb{E}(X^2)-2\mathbb{E}X\cdot\mathbb{E}X+(\mathbb{E}X)^2$

$=\mathbb{E}(X^2)-\mathbb{E}^2(X)$.

e.g. $X\sim P(\lambda)$.　$\mathbb{E}X=\lambda$.

$$\mathbb{D}X=\mathbb{E}(X^2)-\lambda^2=\mathbb{E}(X^2-X)+\mathbb{E}(X)-\lambda^2=\sum_{k=0}^{+\infty}\frac{\lambda^k e^{-\lambda}}{(k-2)!}+\lambda-\lambda^2=\lambda^2\sum_{k=0}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!}e^{-\lambda}+\lambda-\lambda^2$$
$$=\lambda^2+\lambda-\lambda^2=\lambda.$$

e.g. $X\sim N(\mu,\sigma^2)$　$\mathbb{E}X=\mu$.

$\mathbb{D}X=\mathbb{E}(X^2)-\mathbb{E}^2(X)=\sigma^2$

e.g. 几何分布. $\mathbb{E}X=\frac{1}{p}$.　$\mathbb{D}X=\mathbb{E}(X^2-X)+\frac{1}{p}-\frac{1}{p^2}$

$P(X=k)=p(1-p)^{k-1}$.

$=\cdots$

$=\frac{(1-p)}{p^2}$

方差性质：$\mathbb{D}const=0$

$\mathbb{D}(X+const)=\mathbb{D}X$

$\mathbb{D}(aX)=a^2\mathbb{D}X$.

$$\mathbb{D}(X\pm Y)=\mathbb{E}\left((X\pm Y-(\mathbb{E}X\pm\mathbb{E}Y))^2\right)$$
$$=\mathbb{E}\left((X-\mathbb{E}X)^2+(Y-\mathbb{E}Y)^2\pm 2(X-\mathbb{E}X)(Y-\mathbb{E}Y)\right)$$
$$=\mathbb{D}X+\mathbb{D}Y\pm 2\,Cov(X,Y)$$

协方差

$Cov(X,Y)$

[证明] $Cov(X,Y) := \mathbb{E}((X-\mathbb{E}X)(Y-\mathbb{E}Y)) = \mathbb{E}(XY - Y\mathbb{E}X - X\mathbb{E}Y + \mathbb{E}X\cdot\mathbb{E}Y)$

$$= \mathbb{E}(XY) - \mathbb{E}X\,\mathbb{E}Y - \mathbb{E}Y\,\mathbb{E}X + \mathbb{E}X\,\mathbb{E}Y = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y)$$

因此. $X,Y$ 相互独立 $\Rightarrow Cov(X,Y)=0 \Rightarrow \mathbb{D}(X+Y) = \mathbb{D}X + \mathbb{D}Y = \mathbb{D}(X-Y)$.

$\mathbb{D}(X) \leqslant \mathbb{E}((X-c)^2)$. $\forall c$. 当且仅当 $c = \mathbb{E}X$ 取等.  [证]. $\mathbb{E}((X-c)^2) = \mathbb{E}((X-\mathbb{E}X+\mathbb{E}X-c)^2)$. 即可.

~~[证]. $(X-c)^2 = (X-\mathbb{E}X)^2 +$~~

常用期望与方差

| | | | $\mathbb{E}X$ | $\mathbb{D}X$ |
|---|---|---|---|---|
| 0-1分布. | $P(X=1)=p$. | | $p$ | $p(1-p)$ |
| $B(n,p)$ | $P(X=k)=C_n^k p^k (1-p)^{n-k}$ | $k=1,2,\cdots,n$ | $np$ | $np(1-p)$ |
| $P(\lambda)$ | $P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$ | $k=0,1,2,\cdots$ | $\lambda$ | $\lambda$ |
| 几何分布 | $P(X=k)=p(1-p)^{k-1}$ | $0<p<1$, $k=0,1,2\cdots$ | $\frac{1}{p}$ | $\frac{1-p}{p^2}$ |
| $U(a,b)$ | / | | $\frac{a+b}{2}$ | $\frac{(b-a)^2}{12}$ |
| $E(\lambda)$ | $f(x)=\lambda e^{-\lambda x}$ $(x>0)$ | | $\frac{1}{\lambda}$ | $\frac{1}{\lambda^2}$ |
| $N(\mu,\sigma^2)$ | $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ | | $\mu$ | $\sigma^2$ |

标准化随机变量: $D(X)>0$. 称 $X^* = \frac{X-E(X)}{\sqrt{D(X)}}$. $\longrightarrow$ 无量纲数. 可用于平行对比

- 协方差

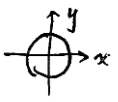
$$Cov(X,Y) := \mathbb{E}((X-\mathbb{E}X)(Y-\mathbb{E}Y)) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y) = \pm\frac{1}{2}\left[\mathbb{D}(X\pm Y) - \mathbb{D}(X) - \mathbb{D}(Y)\right]$$

相关系数: 去除量纲 (标准化).

$$\rho_{XY} = \rho(X,Y) := \mathbb{E}\left(\frac{(X-\mathbb{E}X)(Y-\mathbb{E}Y)}{\sqrt{\mathbb{D}X}\ \sqrt{\mathbb{D}Y}}\right) = Cov(X^*, Y^*) = \mathbb{E}(X^* Y^*).$$

$$= \frac{cov(X,Y)}{\sqrt{\mathbb{D}X}\sqrt{\mathbb{D}Y}}$$

$\rho_{XY}=0$: 不相关 (不线性相关)  $|\rho_{XY}|=1$: 完全相关. (完全线性相关)  $0<|\rho_{XY}|<1$: 部分线性相关

e.g.  不相互独立

$\mathbb{D}(X\pm Y) = \mathbb{D}X \pm \mathbb{D}Y \iff \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) \iff Cov(X,Y)=0 \iff X,Y$不相关

e.g. $(X,Y) \sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho) \Rightarrow cov(X,Y) = \rho\sigma_1\sigma_2$

$\rho=0 \Rightarrow X,Y$ 相互独立 且 不相关.　　或者线性关系. 或者无线性关系. <u>没有部分相关情况</u>

协方差性质 $cov(aX, bY) = ab\, cov(X,Y)$

$cov(X+Y, Z) = cov(X,Z) + cov(Y,Z)$

$cov(X,X) = D(X)$

$|cov(X,Y)|^2 \leq D(X)D(Y)$　　$\boxed{\text{Cauchy-Schwarz不等式}}$

(当且仅当 $\exists t_0$ s.t. $P(Y-E(Y)=t_0(X-E(X)))=1$. 即 $X,Y$ 完全相关 即 $|\rho_{XY}|=1$).

$\Rightarrow 0 \leq |\rho_{XY}| \leq 1$

协方差矩阵

$$C = \begin{pmatrix} cov(x_1,x_1) & cov(x_1,x_2) & \cdots & cov(x_1,x_n) \\ cov(x_2,x_1) & cov(x_2,x_2) & \cdots & cov(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(x_n,x_1) & cov(x_n,x_2) & \cdots & cov(x_n,x_n) \end{pmatrix}$$ where $cov(x_i,x_i) = \mathbb{D}x_i$

e.g. $n$维正态分布

$\boldsymbol{x} = (x_1 \cdots x_n)^T$　　协方差矩阵$\Sigma$.　　$\boldsymbol{x} \sim N(\mathbb{E}\boldsymbol{x}, \Sigma)$.

$\mathbb{E}\boldsymbol{x} = (\mathbb{E}x_1 \cdots \mathbb{E}x_n)^T$

$$f(\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi})^n |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\boldsymbol{x}-\mathbb{E}\boldsymbol{x})^T \Sigma^{-1} (\boldsymbol{x}-\mathbb{E}\boldsymbol{x})}.$$

# Chebyshev 不等式与大数定理

- Chebyshev 不等式

$\mathbb{E}X=\mu$. $\mathbb{D}X=\sigma^2$. 则 $\forall\varepsilon>0$, $P(|X-\mu|\geq\varepsilon)\leq\frac{\sigma^2}{\varepsilon^2}$

$P(|X-\mu|<\varepsilon)>1-\frac{\sigma^2}{\varepsilon^2}$

[证明] $P(|X-\mu|\geq\varepsilon)=\int_{|x-\mu|\geq\varepsilon}1\cdot f(x)\,dx\leq\int_{|x-\mu|\geq\varepsilon}\left|\frac{x-\mu}{\varepsilon}\right|^2 f(x)\,dx=\frac{1}{\varepsilon^2}\int_{|x-\mu|\geq\varepsilon}(x-\mu)^2 f(x)\,dx$

$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2 f(x)\,dx=\frac{\sigma^2}{\varepsilon^2}$. □

e.g. $X,Y$ 在 $[0,1]$ 中随机取两个数. 估计 $P(|\min\{X,Y\}-\mathbb{E}\min\{X,Y\}|<\frac{1}{3})$ 并求其真值.

$f(x,y)=\begin{cases}1, & (x,y)\in[0,1]\times[0,1]\\ 0, & \text{otherwise}\end{cases}$

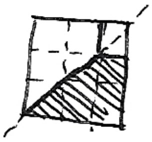
$\mathbb{E}\min\{X,Y\}=\iint_{[0,1]^2}\min\{x,y\}\,dxdy=2\int_0^1\int_0^x y\,dydx=\frac{1}{3}$.

$\mathbb{E}\min^2\{X,Y\}=\cdots=\frac{1}{6}$. $\mathbb{D}\min\{X,Y\}=\frac{1}{18}$.

估计: $P^*=\frac{\frac{1}{18}}{\frac{1}{9}}=\frac{1}{2}$

真值: $P(0<\min\{X,Y\}<\frac{2}{3})=\iint_{[0,1]^2}\min\{x,y\}\,dxdy$

$=2\int_0^{\frac{2}{3}}dx\int_0^x dy+2\int_{\frac{2}{3}}^1 dx\int_0^{\frac{2}{3}}dy=\frac{8}{9}$.



□

- 依概率收敛.

设 $Y_1,Y_2,\cdots Y_n\cdots$ 是一系列随机变量, $X$ 是一个随机变量. 对 $\forall\varepsilon>0$. 有 $\lim_{n\to\infty}P(|Y_n-X|\geq\varepsilon)=0$. 称~.

记作 $Y_n\xrightarrow[n\to\infty]{P}X$. ($X$ 可以是常数)

- Bernoulli 大数定律

$n_A$ 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 相概率.

则有 $Y_n=\frac{n_A}{n}\xrightarrow[n\to\infty]{P}p$. (可用 Chebyshev 证明)

$\boxed{\text{随机变量序列服从大数定律}}$: 若 $X_1,X_2,\cdots X_n,\cdots$ 具有以下性质, 则称.

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{k=1}^n X_k-\frac{1}{n}\sum_{k=1}^n(\mathbb{E}X_k)\right|<\varepsilon\right)=1.$$

- Chebyshev 大数定律:

随机变量序列 $X,\cdots X_n$ 两两不相关, 方差存在且有共同上界. 即

$\mathbb{E}X_k=\mu_k$. $\mathbb{D}X_k=\sigma_k^2\leq\sigma^2$ $(k=1,2,\cdots)$.

则该序列服从大数定律.

"两两不相关"可替为 $\frac{1}{n^2}\mathbb{D}\left(\sum_{k=1}^n X_k\right)\xrightarrow{n\to\infty}0$. (Chebyshev 不等式证)

- Khintchine 大数定律

$X_1, \cdots X_n$ 相互独立且同分布. 若 $E(X_i)=\mu$ 存在.

则 $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{P} \mu$.

[推论] 若k阶原点矩 $(\underbrace{EX_i^k}_{=\mu_k})$ 存在 则 $\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\to\infty]{} \mu_k$

e.g. 可用 $\frac{1}{n}\sum_{i=1}^{n} e^{x_i^2} \xrightarrow[n\to\infty]{} \frac{1}{2}\int_1^3 e^{x^2}dx$ (若 $X_i \sim U(1,3)$). 来求 $\int_1^3 e^{x^2}dx$.

- 中心极限定理

$$\sum_{i=1}^{n} X_i \xrightarrow[\text{以分布收敛于}]{n\to+\infty} N(n\mu; n\sigma^2) \xrightarrow{\text{标准化}} N(0,1)$$

即: $\boxed{\frac{\sum_{i=1}^{n}(X_i - n\mu)}{\sqrt{n\sigma^2}} \xrightarrow[\text{以分布收敛}]{n\to+\infty} N(0,1)}$ 即: $P\left(\frac{\sum_{i=1}^{n}(X_i-n\mu)}{\sqrt{n\sigma^2}} \leq x\right) \longrightarrow \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$

若 $X_1, \cdots X_n$ 是相互独立且同分布. 同时也有 $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{\text{以分布收敛于}}$ 正态分布 $(n\to\infty)$ $N(\mu, \frac{\sigma^2}{n})$

且 $E(X_k)=\mu$, $D(X_k)=\sigma^2$
期望存在 方差存在

标准化后形式一致

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

e.g. 用标准正态分布表示. $X_i \sim E(5)$. $n$ 很大. $P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{5}\right| < \varepsilon\right)$

[解]. $\underbrace{\frac{1}{n}\sum_{k=1}^{n} X_k}_{\frac{1}{5}} \sim N(\underbrace{\tfrac{1}{5}}_{\frac{1}{5}}, \frac{1}{25n})$ $\therefore P(\cdots) = \Phi\left(\frac{\varepsilon}{\sqrt{\frac{1}{25n}}}\right) - \Phi\left(-\frac{\varepsilon}{\sqrt{\frac{1}{25n}}}\right) = 2\Phi(5\varepsilon\sqrt{n}) - 1$ □

$\boxed{\text{DeMoivre - Laplace 中心极限定理}}$

$Y_n \sim B(n,p)$ (即: $Y_n = \sum_{i=1}^{n} X_i$, $X_i = \begin{cases} 1 & A \\ 0 & \bar{A} \end{cases}$. $P(A)=p$) $0<p<1$. 对任意 $x \in$

$$\lim_{n\to\infty} P\left(\frac{Y_n - np}{\sqrt{np(1-p)}} \leq x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

即 $\frac{Y_n - np}{\sqrt{np(1-p)}} \sim N(0,1)$. $Y_n \sim N(np, np(1-p))$

问题: N. 连续分布. 二项分布: 离散分布

$$\underbrace{P(Y_n = x)}_{= C_n^k p^k (1-p)^{n-k}} \cong \varphi(x)\cdot 1 \approx \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \frac{1}{\sqrt{2\pi}\sqrt{np(1-p)}} e^{-\frac{(x-np)^2}{2np(1-p)}} dx$$

↑ 概率密度

$$= \frac{1}{\sqrt{2\pi}\sqrt{np(1-p)}} \exp\left[-(k-np)^2/(2np(1-p))\right]$$

1) $n<10$：直接算

2) $n$较大，$p$较小 $(n \gg \lambda \gg p)$．Poisson分布近似
$(1-p)$
$\lambda = np$适中

3) $n$较大，$0.1<p<0.9$时 或 $n>100$．$p<0.1$ / $p>0.9$：正态分布近似

# 数理统计

总体的一个容量为 $n$ 的

总体. 样本 ← 抽取 $n$ 个个体, 依次观察关心的那个指标 ~ 样本值 观测值 (样本值): $x_1, \cdots x_n$

↓

关心的某一指标的全体

可看作 $n$ 维随机变量 $(X_1, \cdots X_n)$ 的一组可能取值

↑

样本

(总体的一个容量为 $n$ 的样本)

所有 $(X_1 \cdots X_n)$ 取值集合:

第一章样本空间的子集 → 样本空间 $(\mathcal{X})$

(取定了随机事件)

$\boxed{\text{简单随机样本}}$: $(X_1, \cdots X_n)$ 是来自 $X$ 的样本. $X_1, \cdots X_n$ 相互独立 且与 $X$ 同分布.

若 $X: F(X)$. 则 $F(x_1, \cdots x_n) = \prod_{i=1}^{n} F(x_i)$

$\boxed{\text{统计量}}$ $g(x_1, \cdots x_n)$ 是 $n$ 元函数 (无未知参数) ⟶ $g(X_1, \cdots X_n)$ 统计量 <一定是随机变量>

如 $\mu$ 和 $\sigma^2$

给定一个样本值 就能求出值

$g(x_1, \cdots x_n)$ 统计量的 样本观测值.

常用统计量: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 样本均值 vs. $\mathbb{E}X$ (期望/总体均值)

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 样本方差 vs. $\mathbb{D}X$

$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$ 样本标准差

$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ 样本 $k$ 阶原点矩 vs. $\mathbb{E}X^k$

?

$\tilde{S}^2 := CM_k := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$ 样本的 $k$ 阶中心矩 vs. $\mathbb{E}(X - \mathbb{E}X)^2$

顺序统计量: $(X_1, \cdots X_n)$ 样本的样本值为 $(x_1, \cdots x_n)$. 若排序后为 $x_1^* \le x_2^* \le \cdots \le x_n^*$.

定义 随机变量 $X_{(k)} = x_k^*$ (第 $k$ 小的值). $k = 1, 2, \cdots, n$

显然 $X_{(1)} = \min_{1\le k\le n}\{X_k\}$. $X_{(n)} = \max_{1\le k\le n}\{X_k\}$. 称 $X_{(n)} - X_{(1)}$ 为极差.

$$\tilde{X} := \begin{cases} X_{(\frac{n+1}{2})} & n \text{ 为奇} \\ \frac{1}{2}\left(X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}\right) & n \text{ 为偶} \end{cases}$$ 为中位数.

统计量性质: $X_i$ 简单样本. $\mathbb{E}X = \mu$. $\mathbb{D}X = \sigma^2$

$\mathbb{E}(\bar{X}) = \mu$. $\mathbb{D}(\bar{X}) = \frac{\sigma^2}{n}$ (证: $\mathbb{E}\bar{X} = \frac{1}{n}\underbrace{\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)}_{n\mu}$ $\mathbb{D}\bar{X} = \frac{1}{n^2}\underbrace{\mathbb{D}\left(\sum_{i=1}^{n} X_i\right)}_{n\sigma^2} = \frac{\sigma^2}{n}$)

$\mathbb{E}(S^2) = \sigma^2$

[证] $\mathbb{E}S^2 = \mathbb{E}\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) = \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n}(X_i^2 + \bar{X}^2 - 2\bar{X}X_i)\right) = \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n} X_i^2 + n\bar{X}^2 - 2\bar{X}\cdot n\bar{X}\right)$

$= \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right) = \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n} X_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right)^2\right)$

$=\frac{1}{n-1}\sum_{i=1}^{n} E(X_i^2) - \frac{n}{n-1}E(\bar{X}^2) = \frac{1}{n-1}\sum_{i=1}^{n}(DX_i + E^2X_i) - \frac{n}{n-1}(D\bar{X} + E^2\bar{X})$

$=\frac{1}{n-1}(n\sigma^2 + \cancel{n\mu^2}) - \frac{n}{n-1}(\frac{\sigma^2}{n} + \cancel{\mu^2}) = \frac{n-1}{n-1}\sigma^2 = \sigma^2 \longrightarrow$ 因此称 $S^2$ 为样本方差. □

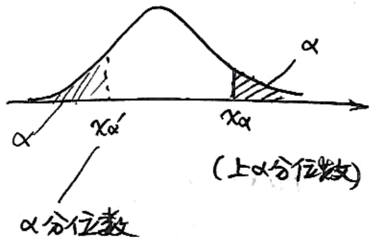
$M_k \xrightarrow[n\to+\infty]{P} E(X^k)$ (若存在)

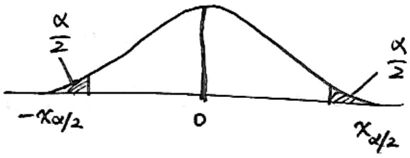
$\boxed{\text{上侧}\alpha\text{分位数}}$: 若 $P(X > x_\alpha) = \alpha$. 称 $x_\alpha$ 为X所服从分布的~.

$\boxed{\alpha\text{分位数}}$: 若 $P(X < x_\alpha) = \alpha$. 称 $x_\alpha$ 为 — ————.

$\boxed{\text{双侧}\alpha\text{分位数}}$: 若X的概率密度函数为偶函数.

$\forall\alpha$. $P(|X| > x_{\alpha/2}) = \alpha$. 称 $x_{\alpha/2}$ 为X所服从分布的双侧α分位数.




标准正态分布的上α分位数

$\Phi(u_{0.95}) = 0.05$.  $-u_\alpha = u_{1-\alpha}$

$P(|X| > u_\alpha) = \min\{2\alpha, 1\}$.

- 抽样分布

$\boxed{\chi^2\text{分布}}$ $X_1, \cdots X_n$ 独立且与X同分布. 都服从 $N(0,1)$

(卡方分布) 称 $\sum_{i=1}^{n} X_i^2$ 服从的分布为自由度n的 $\chi^2$ 分布. 即 $\sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$

$$\chi^2:\quad f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}, & x > 0 \\ 0 & x \le 0 \end{cases}$$

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t}\, dt.$$

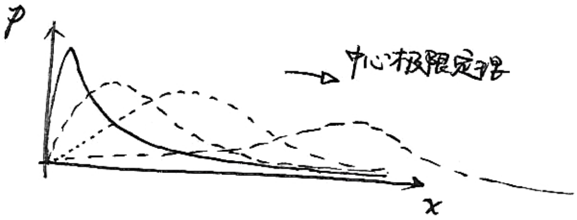
$X_i \sim N(\mu, \sigma) \Rightarrow \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$.

简单样本.

$E(\chi^2(n)) = n$.  $D(\chi^2(n)) = 2n$ ;  $X_1 \sim \chi_1^2(n_1)$, $X_2 \sim \chi_2^2(n_2)$. 相互独立.

$n(\underset{1}{DX_i} + \underset{0}{E^2X_i})$   求 $EX_i^4 - E^2X_i^2$ (用正态分布算)   $X_1 + X_2 \sim \chi^2(n_1 + n_2)$.

$\chi_{1-\alpha} = -\chi_\alpha$




$\boxed{t\text{分布}}$ $X \sim N(0,1)$. $Y \sim \chi^2(n)$  X, Y相互独立

$T = \frac{X}{\sqrt{\frac{Y}{n}}}$   $T \sim T(n)$

自由度为n的T分布

$f_n(t)$ 是偶函数. $\xrightarrow{n\to\infty} \phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

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

$\boxed{F\text{分布}}$ $X\sim\chi^2(m)$. $Y\sim\chi^2(n)$ $X,Y$ 相互独立. $F=\dfrac{X/m}{Y/n}$.

记 $F\sim F(m,n)$.

$n\to\infty$. $\underline{X\to 1.}$ (非分布)

性质: $\dfrac{1}{F}\sim F(n,m)$.

$F_\alpha=\dfrac{1}{F_{1-\alpha}}$ (可用分布函数来证明)

证明: $F_\alpha(1,n)=\left[t_{1-\alpha/2}(n)\right]^2$

[证] $X\sim N(0,1)$. $Y\sim\chi^2(n)$. $T=\dfrac{X}{\sqrt{Y/n}}$. $T^2=\dfrac{X^2}{Y/n}\sim F(1,n)$ 自由度为1的$\chi^2$分布

$\alpha=P(|T|>t_{\alpha/2}(n))$

$=P(T^2>t^2_{\alpha/2}(n))=P(F>t^2_{\alpha/2}(n))$

而 $t_{\alpha/2}=t_{1-\alpha/2}$ $\therefore F_\alpha(1,n)=[t^2_{\alpha/2}(n)]=[t_{1-\alpha/2}(n)]^2$ □

$F\sim F(m,m)$. $P(F>1)=\dfrac{1}{2}$.

$\boxed{\text{正态总体的抽样分布}}$

设 $X\sim N(\mu,\sigma^2)$. $(X_1,\cdots X_n)$是X的简单样本, 则

$\dfrac{\bar X-\mu}{\sigma/\sqrt n}\sim N(0,1)$. 即 $\bar X\sim N(\mu,\dfrac{\sigma^2}{n})$

$\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$ 即 $\dfrac{1}{\sigma^2}\sum\limits_{i=1}^n(X_i-\bar X)^2\sim\chi^2(n-1)$. 感性理解: 有线性关系 $\sum\limits_{i=1}^n\dfrac{X_i-\bar X}{\sigma}=0$ 自由度下降1.

$\dfrac{(n-1)S^2}{\sigma^2}$ 与 $\bar X$ 相互独立

[证] $\sum\limits_i\left(\dfrac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$. ($\because \dfrac{X_i-\mu}{\sigma}$ 相互独立)

[引理] $X_1,\cdots X_n$ 相互独立且 $X_i\sim N(0,1)$. $i=1,2,\cdots,n$. 矩阵C是正交矩阵. ($CC^T=C^TC=E$).

$(Y_1,\cdots Y_n)^T=C(X_1,\cdots X_n)^T$. 则 (1) $Y_1,\cdots Y_n$ 相互独立且 $Y_i\sim N(0,1)$

(2) $Y_i$ 与 $\sum\limits_{j\neq i}Y_j^2$ 相互独立.

$h(y)=Cx$

[证] $(X_1,\cdots X_n)^T=C^{-1}(Y_1,\cdots Y_n)^T$. $J=\left|\dfrac{\partial(x_1,\cdots x_n)}{\partial(y_1,\cdots y_n)}\right|=|C^{-1}|=\mp 1=:|h'(y)|$.

Jacobi行列式

$f_X(x_1,\cdots x_n)=\dfrac{1}{(\sqrt{2\pi})^n}e^{-\frac12 x^Tx}$. $y^Ty=x^T\underset{\overset{\parallel}{E}}{C^TC}x=x^Tx$.

$f_Y(y_1,\cdots y_n)=f_X(C^{-1}(y_1,\cdots y_n)^T)\cdot\underset{\overset{\parallel}{1}}{|h'(y)|}=f_X(C^{-1}(y_1,\cdots y_n)^T)=\dfrac{1}{(\sqrt{2\pi})^n}e^{-\frac12 y^Ty}$

$$\cdots\cdots = (f_{Y_i}(y))^n \longrightarrow Y_1,\cdots Y_n \text{ 相互独立. 且 } Y_i \sim N(0,1)$$

求 $f_{Y_i}$

令 $Y_i = \dfrac{X_i-\mu}{\sigma} \sim N(0,1)$ $\quad \bar{Y} = \frac{1}{n}\sum_i Y_i \sim N(0,n) \xrightarrow{\text{标准化}} \dfrac{Y_1+\cdots+Y_n}{\sqrt{n}} = \sqrt{n}\,\bar{Y} = \sqrt{n}\,\dfrac{\bar{X}-\mu}{\sigma}$.

构造正交变换 $Z = \begin{pmatrix} Z_1 \\ \vdots \\ Z_n \end{pmatrix} = \tilde{C}\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$. $\quad \tilde{C} = \begin{pmatrix} \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ C_{21} & \cdots & C_{2n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nn} \end{pmatrix}$ s.t. $\tilde{C}^T\tilde{C} = \tilde{C}\tilde{C}^T = E$

则 $Z_i \sim N(0,1)$ 且 i.i.d.

$$Z_1 = \frac{Y_1+\cdots+Y_n}{\sqrt{n}} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

$$Z_2^2+\cdots+Z_n^2 = (Z_1^2+\cdots+Z_n^2) - Z_1^2 = z^Tz - Z_1^2 = y^Ty - Z_1^2 = Y_1^2+\cdots+Y_n^2 - \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

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

(左侧是$(n-1)$个服从$N(0,1)$的随机变量)

$Z_1, Z_2, \cdots Z_n$ 相互独立 $\Rightarrow \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ 与 $\dfrac{(n-1)S^2}{\sigma^2}$ 相互独立 $\Rightarrow \bar{X}$ 与 $\dfrac{(n-1)S^2}{\sigma^2}$ 相互独立 $\square$

考虑到 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$. $\Rightarrow$ $t$分布: $\dfrac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{(n-1)\sigma^2}}} = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

因此. $X$ 总体方差已知 $\Rightarrow$ 用 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

未知 $\Rightarrow$ 用 $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

$S$: 样本标准差.

对两个正态总体的抽样分布.

$X$、$Y$ 独立. 取$(X_1,\cdots X_n)$. $(Y_1\cdots Y_m)$ 两个样本(简单样本).

$X \sim N(\mu_1,\sigma_1)$, $Y \sim N(\mu_2,\sigma_2)$

$\bar{X}$, $S_1^2$, $\bar{Y}$, $S_2^2$ 彼此相互独立

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} \sim N(0,1). \qquad \frac{(n-1)S_1^2}{\sigma_1^2} + \frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2(n+m-2)$$

当 $\sigma_1 = \sigma_2 = \sigma$ 时. 可知

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} \Big/ \sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{(n+m-2)\sigma^2}} = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}} \Big/ \sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}$$

$$\sim t(n+m-2)$$

# 参数估计

- 频率替代法（点估计）

$\hat{\theta}$：估计值 ⟶ 选不同参照有不同估计值.

- 矩法估计（点估计）

用样本k阶矩作为总体k阶矩估计量. ⟵ Khintchine 大数定律推论

e.g. $\mathbb{E}(X) = \mu \approx \bar{X} =: \hat{\mu}$

$\mathbb{E}(X^2) = \mu^2 + \sigma^2 \approx \frac{1}{n}\sum_i X_i^2 \Rightarrow \begin{cases} \hat{\mu} = \bar{X} \\ \hat{\sigma}^2 = \frac{1}{n}\sum_i (X_i - \bar{X})^2 = \frac{1}{n}\sum_i X_i^2 - n\bar{X}^2 \end{cases}$

普遍形式：

$$\begin{aligned} \mathbb{E}(X) &= \mu_1(\theta_1, \cdots \theta_k) \approx \frac{1}{n}\sum_i X_i = \bar{X} \\ \mathbb{E}(X^2) &= \mu_2(\theta_1, \cdots \theta_k) \approx \frac{1}{n}\sum_i X_i^2 \\ &\vdots \\ \mathbb{E}(X^k) &= \mu_k(\theta_1, \cdots \theta_k) \approx \frac{1}{n}\sum_i X_i^k \end{aligned} \Rightarrow \begin{cases} \hat{\theta}_1(X_1, \cdots X_n) \\ \vdots \\ \hat{\theta}_k(X_1, \cdots X_n) \end{cases}$$

矩估计量

↑ 随机变量

$\xRightarrow{代入 x_1 \cdots x_n} \hat{\theta}_1, \cdots \hat{\theta}_k$ 矩估计值

问题：1) 不同分布对应的k阶矩估计值一样； 2) 可能不唯一. e.g. $\begin{cases} \hat{\lambda} = \frac{1}{\bar{X}} \\ \hat{\lambda} = \sqrt{1/CM_2} \end{cases}$ ⟵ 一般选低阶

- 点估计的最大似然估计法

使一次试验就出现的事件有较大概率的参数作为估计.

即使 $P(X_1 = x_1, X_2 = x_2, \cdots X_n = x_n)$ 达到最大的参数

$= \cdots = L(\theta, x) \Rightarrow$ 即 $\ln L(\theta)$ 最大 ⟶ $\hat{\theta}_1 \cdots \hat{\theta}_k$：最大似然估计值

Likelihood.

最大似然函数

离散型：$L = \prod_i P(x_i, \theta_1, \cdots \theta_k)$. 连续型：$L = \prod_i f(x_i, \theta_1, \cdots \theta_k)$

最大似然估计的不变性：设 $\hat{\theta}$ 是 $\theta$ 的最大似然估计值

若 $u(\theta)$ 是 $\theta$ 的函数，且具有单值反函数 $\theta = \theta(u)$.

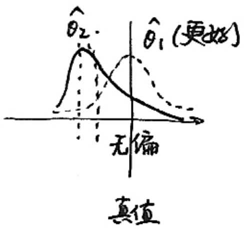
则 $u(\hat{\theta})$ 是 $u(\theta)$ 的最大似然估计值. 对应形式为 …… 量.

V.S. 矩法估计：不具有不变性

• 评估标准

**无偏性** 记 $(X_1, \cdots X_n)$ 为总体 $X$ 的样本，$\theta$ 是总体 $X$ 的一个参数. $\hat{\theta} = \hat{\theta}(X_1, \cdots X_n)$ 是 $\theta$ 的估计量.

若 $\mathbb{E}(\hat{\theta})$ 存在且 $\mathbb{E}(\hat{\theta}) = \theta$. 称 $\hat{\theta}$ 具有. 即 $\hat{\theta}$ 是 $\theta$ 的无偏估计量.



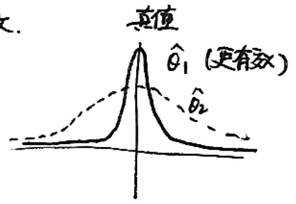

结论：$\mu_k = \mathbb{E}(X^k)$. $\mu = \mathbb{E}(X)$. $\sigma^2 = \mathbb{D}(X)$. 简单样本.

1) $\bar{X}$ 是 $\mu$ 的无偏估计量；2) $M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ 是 $\mu_k$ 的无偏估计量；

3) $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 是 $\sigma^2$ 的无偏估计量.

无偏校正：e.g. 最大似然. $\mathbb{E}(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2 \longrightarrow$ 校正为 $\frac{n}{n-1}\hat{\sigma}^2$. 本质：期望的线性性

**有效性** 若 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是无偏估计量且 $\mathbb{D}(\hat{\theta}_1) < \mathbb{D}(\hat{\theta}_2)$. 则 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效.




**相合性(一致性)** 设 $\hat{\theta} = \hat{\theta}(X_1, \cdots X_n)$ 是 $\theta$ 的估计量.

若 $\hat{\theta} \xrightarrow{P} \theta$. 称 $\hat{\theta}$ 为 $\theta$ 的一致性估计量.

若 $\hat{\theta}$ 是 $\theta$ 的无偏估计量且 $\lim_{n\to\infty} \mathbb{D}(\hat{\theta}) = 0$. 则 $\hat{\theta}$ 是 $\theta$ 的一致性估计量.

**Rao-Cramer 不等式** $\hat{\theta}$ 是 $\theta$ 的无偏估计量.

$$\text{则 } \mathbb{D}(\hat{\theta}) \geq \frac{1}{n\mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln p(X,\theta)\right)^2\right]} =: \mathbb{I}(\theta)$$

概率分布 $\begin{cases} \text{离散：分布列} \\ \text{连续：概率密度} \end{cases}$

若存在 $\hat{\theta}$ 使 $\mathbb{D}(\hat{\theta}) = \mathbb{I}(\theta)$. 称 $\hat{\theta}$ 为(最)有效估计

▷ 区间估计

找一个能使包含待估参数的概率为一定值的区间.
(含有)　　　置信区间　　　置信度(可信度)

注：落在 vs 含有
该值(落的)是随机　　区间是随机

正态总体区间估计 —— 常用：$\Phi(1.96) = 0.9750$ 即 $\boxed{u_{0.025} = 1.96}$

• 含有平均值 $\mu$ 的概率达到 $(1-\alpha)$ 的区间：$\left(\bar{X} - \frac{\sigma_0}{\sqrt{n}}u_{\alpha/2},\ \bar{X} + \frac{\sigma_0}{\sqrt{n}}u_{\alpha/2}\right)$ —— 均值的置信区间

$$P\left(\left|\frac{\bar{X}-\mu}{\sigma_0/\sqrt{n}}\right| < u_{\alpha/2}\right) = 1-\alpha \Rightarrow \cdots\cdots$$

枢轴量

注：待估参数 $\theta$ 的有效估计不一定存在！

① 含待估参数且不含其它未知参数
② 服从已知分布

使用 $(-u_{\frac{\alpha}{2}}, u_{\frac{\alpha}{2}})$ 而非 $(-u_{\alpha/3}, u_{2\alpha/3})$. ? $\longrightarrow$ $(-u_{\frac{\alpha}{2}}, u_{\frac{\alpha}{2}})$ 的区间长度最短

(是最短的置信区间)

$g(\theta, X_1, \cdots X_n)$ ← 枢轴量

置信区间: $\theta$: 待估参数  $X_1, \cdots X_n$ 为样本. 构造 $\hat\theta_1 . \hat\theta_2$  $\Theta = (\hat\theta_1, \hat\theta_2)$ 使 $P\left(g(\hat\theta_1, X_1, \cdots X_n) \le g(\theta, X_1, \cdots X_n) < g(\hat\theta_2, X_1, \cdots X_n)\right)$

~~使 $P(\hat\theta_1 < g(\theta, X_1 \cdots X_n) < \hat\theta_2)$~~  称 $\Theta$.

一般在确定 $\alpha$ 后选最小置信区间. (提高精度 $\hat\theta_2 - \hat\theta_1$)

枢轴量: 一般选取有效估计量. 若不存在选近优.

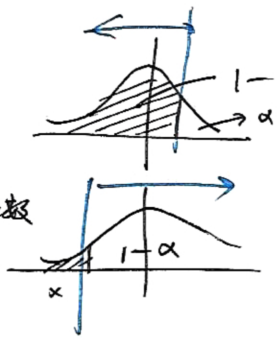
$\sigma^2$ 置信区间: 枢轴量 $\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \begin{matrix}\mu \\ \bar X\end{matrix}\right)^2$ —— 取决于 $\mu$ 是否已知。 $\longrightarrow$ 一般用 $\frac{\alpha}{2} . \frac{\alpha}{2}$

($n \to \infty$. $\to$ 对称)

- 单侧置信区间

$\frac{1-\alpha.}{}$ (置信度) $\longrightarrow$ 上 $\alpha$ 分位数

或上 $(1-\alpha)$ 分位数




(只要单侧)

上限  $u_\alpha$ / $t_\alpha(n-1)$ / $\cdots$

下限

- 两个正态总体的置信区间

一般要求相互独立. 置信度 $1-\alpha$.

① $\sigma_1^2 . \sigma_2^2$ 已知.  枢轴量 $\dfrac{\bar X - \bar Y - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim N(0,1)$

② $\sigma_1^2 . \sigma_2^2$ 未知. $\sigma_1^2 = \sigma_2^2 = \sigma$.  枢轴量 $\dfrac{\bar X - \bar Y - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}} \sim t(n+m-2)$

$\frac{(n-1)S_1^2}{\sigma^2} + \frac{(m-1)S_2^2}{\sigma^2} \sim \chi^2(n+m-2)$

$\dfrac{N}{\sqrt{\frac{\chi^2}{N(total)}}}$

③ $\sigma_1^2 . \sigma_2^2$ 未知. $n, m > 50$. 近似正态

枢轴量 $\dfrac{\bar X - \bar Y - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}} \sim N(0,1)$

④ $n = m$. $\sigma_1^2 . \sigma_2^2$ 未知.  ~~枢轴量~~  $Z_i = X_i - Y_i$

↓

新的总体: $X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$

枢轴量 $\dfrac{\bar z - (\mu_1 - \mu_2)}{S/\sqrt{n}} \sim t(n-1)$

$S = \sum_i (Z_i - \bar z)^2 / (n-1)$

⑤ $\mu_1, \mu_2$ 未知. $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信区间

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

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

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

$$F = \frac{\chi^2/n}{\chi^2/n}$$

- 非正态总体

中心极限定理. $n \to \infty$. $\longrightarrow N$

# 假设检验

原假设(原设)、 备择假设

$\mu=\mu_0$ 记作 $H_0$　　　$\mu\neq\mu_0$ 记作 $H_1$.

检验统计量: $\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ 若 $X\sim N(\mu,\sigma)$

如何认定原假设是否正确? —— 分位数(一个统一标准)

$P\left(\left|\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| > u_{\alpha/2}\right)=\alpha \Rightarrow \left|\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| > u_{\alpha/2}$ : $\boxed{\text{拒绝域}}$ → 拒绝 $H_0$.

小概率原理: 认为小概率事件基本不发生.

$\alpha$: "标准" $\boxed{\text{检验水平(显著性水平)}}$

$\leq u_{\alpha/2}$ : 接受域

"符合" ~ 双侧检验　　　"不小于" ~ 单侧检验

$\bar{X}\sim N(\mu,\sigma)$

$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ ← 未知

$\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$　　$\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$　　$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

$\left\{\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < -u_\alpha\right\} \subset \left\{\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} < -u_\alpha\right\}$

① 假设 ② 检验统计量 ③ 区间 ④ 接受/拒绝

| | 认为 $H_0$ 真 | 认为 $H_0$ 假 |
|---|---|---|
| $H_0$ 真 | ✓ | 弃真(概率: $\alpha$ ("$\mu=\mu_0$" =: $H_0$); $\leq\alpha$ ("$\mu\leq\mu_0$" =: $H_0$)) |
| $H_0$ 假 | 纳伪 ($\beta$) | ✓ ($1-\beta$) |

$\alpha\downarrow$. 弃真$\downarrow$ 纳伪$\uparrow$

$n\uparrow$　$\downarrow$　$\downarrow$

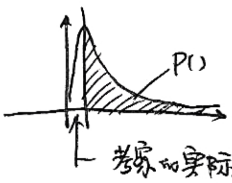
目的: 让两类错误的概率都尽可能小

p值检验法

① U检验　　样本 → $u_0$　　算 $P(|\ \ |>u_0)=\square$　　$>\alpha$: 接受　$<\alpha$: 拒绝

② $\chi^2/t$ 检验: $P(|\text{检验统计量}| > t_0 \text{ 或 } \chi_0^2) = 2P(\text{检}\sim > t_0/\chi_0^2)$

若在左极端. $2(1-P(\ \ ))$




考察的实际是大小的拒绝域.

# 一些期末复习

- 古典概型　　　　排列组合题)

概率公理化定义（虽然感觉不会考到）：

E是一系列随机试验，$\Omega$ 为其对应样本空间．若存在映射 $P:\Omega\to\mathbb{R}$ s.t.

(1) [redacted]，$\forall A\subseteq\Omega$．$P(A)\geq 0$　　(2) $P(\Omega)=1$．

(3) 对两两互不相容事件 $A_1, A_2, A_3, \cdots A_n, \cdots$　　有 $P(\bigcup_{i=1}^{\infty} A_i)=\sum_{i=1}^{\infty} P(A_i)$

称 P(A).

容斥原理．略．　$P(\bigcup_{i=1}^{n} A_i)=\sum_{l=1}^{n}\sum_{1\leq k_1<\cdots<k_l\leq n} P(\bigcup_{i=1}^{l} A_{k_i})\,(-1)^{l+1}$

- 条件概率与 Bayes 公式与事件独立性

$$P(B|A)=\frac{P(AB)}{P(A)}\quad (P(A)>0)$$

A条件下B发生

$\Longleftrightarrow$ 乘法公式：$P(AB)=P(A)P(B|A)=P(B)P(A|B)$　$(P(A)>0,\ P(B)>0)$

全概率公式：记 $B_1,\cdots B_n$ 为 $\Omega$ 的一个划分．即 $\forall\{i,j\}\subseteq\{1,\cdots,n\}$ $B_iB_j=\emptyset$．且 $\bigcup_{i=1}^{n}B_i=\Omega$．

对任意事件A，有 $P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$　$(P(B_i)>0)$

$\Updownarrow$

Bayes 公式

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

* P(A)：先验概率．　P(A|B)：后验概率

事件独立性：$P(AB)=P(A)P(B)\Longleftrightarrow P(A|\bar{B})=P(A|B)=P(A)$ [redacted]

[redacted]

~~随机变量与期望与分布与方差~~

~~连续型/离散型．略~~

[redacted]

- 随机变量与分布与期望与方差

连续型/离散型．略．

从已知概率分布推相关事件概率分布：已知 $f_X(x)$．$g(x)$ 在 $\mathbb{R}$ 上严格单调可导．$Y=g(X)$

则
$$f_Y(y)=\begin{cases} f_X(g^{-1}(y))\,\left|(g^{-1})'(y)\right| & \alpha<y<\beta \\ 0 & \text{otherwise}\end{cases}$$

导数！

$\{\alpha,\beta\}=\{g(-\infty),g(+\infty)\}$

二维随机变量

联合分布 $F(X,Y)$．联合概率密度 $f_{X,Y}(x,y)$

边缘分布 $F_X(x)=P(X\leq x)=F(x,+\infty)=\lim_{y\to+\infty}F(x,y)$

(画画图就显然了)

* $F(+\infty,+\infty)=1$．　$F(-\infty,-\infty)=F(-\infty,+\infty)=F(+\infty,-\infty)=0$．

边缘概率密度：$f_X(x)=\int_{-\infty}^{+\infty}f_{X,Y}(x,y)\,dy$

[redacted]．条件分布：$f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}=\dfrac{f(x,y)}{\int_{-\infty}^{+\infty}f(x,y)\,dy}$

固定值　　$>0$

二维：从已知概率分布推相关事件概率分布 $\longrightarrow$ 从分布函数下手.

$\mathbb{E}g(X)=\int_{-\infty}^{+\infty} g(x)f_X(x)\,dx$ （以一维连续为例）

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

$\operatorname{cov}(X,Y)=\mathbb{E}\left((X-\mathbb{E}X)(Y-\mathbb{E}Y)\right)$

$\rho_{XY}=\dfrac{\operatorname{cov}(X,Y)}{\sqrt{\mathbb{D}X}\sqrt{\mathbb{D}Y}}=\mathbb{E}\dfrac{(X-\mathbb{E}X)(Y-\mathbb{E}Y)}{\sqrt{\mathbb{D}X}\sqrt{\mathbb{D}Y}}$

期望 : $\mathbb{E}(aX)=a\mathbb{E}X$ $\mathbb{E}(X+Y)=\mathbb{E}X+\mathbb{E}Y$

方差 : $\mathbb{D}(aX)=a^2\mathbb{D}X$

协方差 : $\operatorname{cov}(aX,bY)=ab\operatorname{cov}(X,Y)$

相关系数

期望: <u>相互独立时</u> $\mathbb{E}(XY)=(\mathbb{E}X)(\mathbb{E}Y)$

始终有 $\mathbb{D}X=\mathbb{E}(X^2)-(\mathbb{E}X)^2$

方差: $\mathbb{D}(X\pm Y)=\mathbb{D}(X)+\mathbb{D}(Y)+2\operatorname{cov}(X,Y)$

$\operatorname{cov}(X,Y)=\mathbb{E}(XY)-(\mathbb{E}X)(\mathbb{E}Y)$.

不相关. $\operatorname{cov}(X,Y)=0$. 即 $\rho_{XY}=0$ $\not\Rightarrow$ 独立 $\Leftarrow$

*估计用得上: $\mathbb{D}(X)\leq\mathbb{E}((X-c)^2)$

"=": iff. $c=\mathbb{E}X$

常用分布及其期望与方差

| | | | $\mathbb{E}X$ | $\mathbb{D}X$ |
|---|---|---|---|---|
| 离散型 | 0-1分布 | $P(X=1)=p$, $P(X=0)=1-p$, $p\in(0,1)$ | $p$ | $p(1-p)$ |
| | 二项分布 $B(n,p)$ | $P(X=k)=C_n^k p^k(1-p)^{n-k}$, $p\in(0,1)$, $k=0,1,2,\dots,n$ | $np$ | $np(1-p)$ |
| | Poisson分布 $P(\lambda)$ | $P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$, $\lambda>0$, $k=0,1,2,\dots$ | $\lambda$ | $\lambda$ |
| | 几何分布 | $P(X=k)=p(1-p)^{k-1}$, $0<p<1$, $k=0,1,2,\dots$ | $\dfrac{1}{p}$ | $\dfrac{1-p}{p^2}$ |
| 连续型 | 均匀分布 $U(a,b)$ | $f_X(x)=\begin{cases}\frac{1}{b-a} & x\in(a,b)\\ 0 & \text{otherwise}\end{cases}$ | $\dfrac{a+b}{2}$ | $\dfrac{(b-a)^2}{12}$ |
| | 指数分布 $E(\lambda)$ | $f_X(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & \text{otherwise}\end{cases}$, $\lambda>0$ | $\dfrac{1}{\lambda}$ | $\dfrac{1}{\lambda^2}$ |
| | 正态分布 $N(\mu,\sigma)$ | $f_X(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ | $\mu$ | $\sigma^2$ |
| | 卡方分布 $\chi^2(n)$ | / | $n$ | $2n$ |
| | t分布* $t(n)$ | / | 0（显然） | $\dfrac{n}{n-2}$（应该不会考到） |
| | F分布* $F(m,n)$ | / | $\dfrac{n}{n-2}$ | $\dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$（不考了吧） |

$\boxed{\text{Poisson定理}}$ $\lim\limits_{n\to\infty} np_n=\lambda$

对固定k有 $\lim\limits_{n\to\infty} C_n^k p_n^k(1-p_n)^{n-k}=e^{-\lambda}\dfrac{\lambda^k}{k!}$

即二项分布推到极限为Poisson分布.

* 转动惯量（草

线性函数也是正态分布.

二维正态的边缘分布也是正态分布.

$X_i\sim N(0,1)$

$\sum_{i=1}^n X_i^2\longrightarrow\chi^2(n)$

$\dfrac{N(0,1)}{\sqrt{\frac{\chi^2(n)}{n}}}=t(n)$

$\dfrac{\chi^2(m)/m}{\chi^2(n)/n}=F(m,n)$

常用 Trick: 1) 一些积分化为常用分布的~~概率~~/方差/期望./[redacted]/平方期望

e.g. $\int_0^{\infty} \frac{x^2}{\theta} e^{-\frac{x^2}{2\theta^2}} dx = \frac{\sqrt{2\pi}}{2} \int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi}\,\theta} e^{-\frac{x^2}{2\theta^2}} dx = \mathbb{E}(Y^2) = \mathbb{D}(Y) + \mathbb{E}^2(Y)$ $\quad Y \sim N(0, \theta^2)$

$= \theta^2 + 0 = \theta^2$

$\int_0^{\infty} x^2 e^{-\lambda x} dx = \frac{1}{\lambda} \int_0^{\infty} x^2 \underbrace{\lambda e^{-\lambda x}}_{E(\lambda)} dx = \frac{1}{\lambda} \mathbb{E}(Y^2) = \frac{1}{\lambda}\left(\frac{1}{\lambda^2} + \frac{1}{\lambda^2}\right) = \frac{2}{\lambda^3}$ $\quad Y \sim E(\lambda)$

* 注意凑出的可能是平方的期望而非方差!

2) 棘手了就看分布函数./画图

e.g. $F_{Max}(x) = P(Max\{X_1, X_2, \cdots X_n\} \leq x) = F\underbrace{(x, x, \cdots, x)}_{n个}$

$\xlongequal[\text{且同分布 } F_X]{\text{若 } X_1, \cdots X_n \text{ 相互独立}} [F_X(x)]^n$

- Chebyshev 不等式与大数定[redacted]律与中心极限定理

Chebyshev 不等式. $\mathbb{E}X = \mu$. $\mathbb{D}X = \sigma^2$. $\forall \varepsilon > 0$ $\quad P(|X-\mu| < \varepsilon) > 1 - \frac{\sigma^2}{\varepsilon^2}$ (反号)

$\boxed{P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}}$ [redacted]

依概率收敛 $\lim_{n\to\infty} P(|X_n - Y| \geq \varepsilon) = 0$. $\quad X_n \xrightarrow[n\to+\infty]{P} Y$ (随机变量. 可以是常数)

* Bernoulli 大数定律 $\frac{n_A}{n} \xrightarrow[n\to\infty]{P} p_A$. (一频数)

〈感觉没太大用〉

随机变量序列 $\{X_n\}$ 服从大数定律 $\quad \lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}(X_k)\right| < \varepsilon\right) = 1$

Chebyshev 大数定律: $X_1, \cdots X_n$ 两两不相关[1]. 方差存在且有共同上界

即 $\mathbb{E}X_k = \mu_k$. $\quad \mathbb{D}X_k \leq \sigma^2$

则 $\{X_n\}$ 服从大数定律

[1] 可改为

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

可用 Chebyshev 证

Khintchine 大数定律: $X_1, \cdots X_n$ 相互独立且同分布 (i.i.d.)

$\mathbb{E}(X_i) = \mu$. $\quad \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to+\infty]{P} \mu$.

[推论] $\frac{1}{n}\sum_{k=1}^{n} X_k^t \xrightarrow[n\to+\infty]{P} \mu_t := \mathbb{E}(X^t)$ (若存在)

中心极限定理: 万物皆正态

若 $\{X_i\}$ i.i.d. $\mathbb{E}X = \mu$. $\mathbb{D}X = \sigma^2$. $\quad \frac{\sum_{i=1}^{n}(X_i - n\mu)}{\sqrt{n\sigma^2}} \xrightarrow[\text{以分布收敛}]{n\to+\infty} N(0,1)$ 即 $\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n}(X_i - n\mu)}{\sigma\sqrt{n}} \leq x\right)$

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

* $\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$ 标准化

* De Moivre-Laplace 中心极限定理：二项推无穷成正态

$$Y_n \sim B(n,p) = \sum_{i=1}^{n} \underset{\text{0-1分布}}{X_i} \quad 0<p<1. \qquad \frac{Y_n - np}{\sqrt{np(1-p)}} \overset{\text{近似}}{\sim} N(0,1)$$

$$* \; P(Y_n = y) \cong \varphi(y)\cdot 1 \approx \int_{y-\frac{1}{2}}^{y+\frac{1}{2}} \frac{1}{\sqrt{2\pi}\sqrt{np(1-p)}} e^{-\frac{(x-np)^2}{2np(1-p)}} dx$$

① $n<10$ 二项分布直接上

② $n$较大 $n >> \lambda = np >> p$. Poisson分布

③ $n$较大. $<100 \rightarrow 0.1<p<0.9$

$>100 \rightarrow p<0.1$ 或 $p>0.9$ 正态分布.

- 数理统计与参数估计与假设检验

简单随机样本 $X_1, \cdots X_n$ 与 $X$ i.i.d. $\mathbb{E}X = \mu$. $\mathbb{D}X = \sigma^2$. 统计量. 样本观测值. 随机变量

样本均值 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 样本方差

$\mathbb{E}\bar{X} = \mu$.

$\mathbb{E}S^2 = \sigma^2$

* 样本k阶原点矩 $M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$

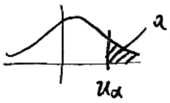
中心 $CM_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$

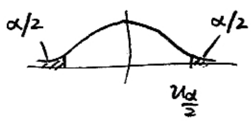
$* \; M_k \xrightarrow[n\to+\infty]{P} \mathbb{E}(X^k)$.

顺序统计量 $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ 且 $\{X_{(1)}, \cdots X_{(n)}\} = \{X_1, \cdots X_n\}$

$X_{(1)} = \min\limits_{1\leq k\leq n}\{X_k\}$, $X_{(n)} = \max\limits_{1\leq k\leq n}\{X_k\}$.

上α分位数 * 双侧α分位数（要求 $f_X(x)$ 偶函数）




$\chi^2$分布. t分布. F分布（上上页）

$X \sim N(\mu, \sigma^2)$. $(X_1, \cdots X_n)$是简单样本. 则

$$\begin{cases} \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ 即 } \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \\ \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \text{ 即 } \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \sim \chi^2(n-1) \\ \bar{X} \text{ 与 } S^2 \text{ 相互独立} \end{cases}$$

点估计

1)* 频率替代法

2) 矩法估计 $\mathbb{E}(X^k) = \mu_k(\theta_1, \cdots \theta_n)$ 则 $\left.\begin{matrix}\bar{X} \; (k=1) \\ M_k \; (k\neq 1)\end{matrix}\right. = \mu_k(\hat{\theta}_1, \cdots \hat{\theta}_n) \Rightarrow \begin{cases}\hat{\theta}_1 = \cdots \\ \vdots \\ \hat{\theta}_n = \cdots\end{cases}$

↑ 一般选 最低阶含θ

3) 最大似然估计.

最大似然函数 $L(\theta) = \prod_{i}^{P} f(x_i, \theta_1, \cdots \theta_n)$ → $\frac{\partial L}{\partial \theta} = 0$ → $\frac{\partial \ln L}{\partial \theta} = 0$ （使L最大）

无偏性：$\mathbb{E}(\hat{\theta}) = \theta$. ← 无偏修正（乘常数使 $\mathbb{E}(\underset{\text{无偏修正后的估计值}}{\underline{k\hat{\theta}}}) = \theta$

有效性：无偏 $\hat{\theta}_1 . \hat{\theta}_2$. $\mathbb{D}(\hat{\theta}_1) \leq \mathbb{D}(\hat{\theta}_2)$. $\hat{\theta}_1$ 更有效

*相合性（一致性）：$\hat{\theta} \xrightarrow[n\to+\infty]{P} \theta$. （对无偏：$\lim\limits_{n\to\infty} \mathbb{D}(\hat{\theta}) = 0$）

| 区间估计 | | | 枢轴量 | 双侧置信区间 | (单侧类似)(可现推) |
| --- | --- | --- | --- | --- | --- |
| | $\mu$. | $\sigma$已知 | $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ | $\left(\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2},\ \bar{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right)$ | $\left(\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha},\ +\infty\right)$ |
| | | $\sigma$未知 | $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ | $\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}^{(n-1)},\ \bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}^{(n-1)}\right)$ | $\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha}^{(n-1)},+\infty\right)$ |
| 置信度$(1-\alpha)$ | $\sigma$ | $\mu$已知 | $\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$ | $\left(\frac{\square}{\chi^2_{\alpha/2}(n)},\ \frac{\square}{\chi^2_{1-\alpha/2}(n)}\right)$ | $\left(\frac{\square}{\chi^2_{\alpha}(n)},\ +\infty\right)$ |
| | | $\mu$未知 | $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ | $\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\ \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$ | $\left(-\infty,\ \frac{\square}{\chi^2_{1-\alpha}(n-1)}\right)$ |

2个正态略(不考对吧)

对$\mu$: 接受域 $\uparrow *$

| 假设检验 | | | 检验统计量 | 拒绝域 | (也可现推) |
| --- | --- | --- | --- | --- | --- |
| | $\mu$ | $\sigma$已知 | $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ | $\left\lvert\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right\rvert \geq u_{\alpha/2}$ | |
| 检验水平 (显著性水平) | | $\sigma$未知 | $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ | $\left\lvert\frac{\bar{X}-\mu}{S/\sqrt{n}}\right\rvert \geq t_{\alpha/2}^{(n-1)}$ | |
| $\alpha$ | $\sigma$ | $\mu$已知 | $\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$ | $\{\square < \chi^2_{1-\alpha/2}(n)\} \cup \{\square > \chi^2_{\alpha/2}(n)\}$ | |
| | | $\mu$未知 | $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ | $\left\{\frac{(n-1)S^2}{\sigma^2} < \chi^2_{1-\alpha/2}(n-1)\right\} \cup \left\{\frac{(n-1)S^2}{\sigma^2} > \chi^2_{\alpha/2}(n)\right\}$ | |

对估计的参数而言

"中心"附近

区间估计 —— $P$(枢轴量 $\in$ 置信区间) $= 1-\alpha \implies$ 推置信区间.

假设检验 ——
1) 零假设与备择假设 $H_0$: … $H_1$: …
2) 检验统计量: $\square \sim ?(\cdots)$
3) 拒绝域: $P$(检验统计量 $\in$ 拒绝域) $= \alpha \implies$ 推拒绝域

边缘的那些

对检验统计量

4) 检验统计量的样本观测值

↳ 判断 接受/拒绝 $H_0$

补: 两个正态分布总体

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

$$\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1, n_2)$$

* 若 $\sigma_1=\sigma_2=\sigma$ 未知

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}} \sim t(n_1+n_2-2)$$

$$\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1, n_2)$$