1 随机事件和概率

随机事件及其运算

基本定义

随机试验具有**可重复性**、**可能结果多样性**、**不可预测性**。

随机试验 E 所有可能的结果组成的集合称为**样本空间**，记作 Ω；样本空间的元素，也就是随机试验 E 的直接结果，称为**样本点**，记作 ω。

在随机试验中可能发生也可能不发生的事件称为**随机事件**，用大写英文字母 A, B, C 记录。随机事件的本质是样本点的集合，也就是 Ω 的子集，因此有：

$$A \subseteq \Omega \Longrightarrow A\text{为随机事件}$$

同时，

$$\text{随机事件}A\text{发生} \Longleftrightarrow \omega \in A$$

随机事件是单点集，则称作**基本事件**；是全集 Ω，称为**必然事件**；是空集 $\varnothing$，称为**不可能事件**。

随机事件的运算及运算律

1. $A \subseteq B \Longleftrightarrow (\forall \omega)(\omega \in A \rightarrow \omega \in B)$，即 A 发生 B 一定发生。
2. $A = B \Longleftrightarrow A \subseteq B \wedge B \subseteq A$，即 A, B 同时发生。
3. 和事件 $A \cup B$，也记作 $A + B$，即 $A \cup B$ 发生，则 A, B 至少其一发生。

   多个事件的和事件 $\cup_{i=1}^{n} A_i$，n 可以为 $+\infty$。
4. 积事件 $A \cap B$，也记作 AB，即 $A \cap B$ 发生，则 A, B 必共同发生。

   多个事件的积事件 $\cap_{i=1}^{n} A_i$，n 可以为 $+\infty$。
5. **互不相容**（互斥事件）若 $AB = \varnothing$，则称 A, B 互不相容（或为互斥事件）。

   若 $A_1, A_2, \ldots, A_n$ 两两互斥，即为 $A_i A_j = \varnothing \ (i, j = 1, 2, \ldots, n)$，$n$ 可以为 $+\infty$。
6. **逆事件**（对立事件）若 $A \cap B = \varnothing, A \cup B = \Omega$，则称 B 是 A 的对立事件（逆事件），记作 $\bar{A} = B$。
7. **差事件** $A - B$，即 A 发生且 B 不发生，即 $A\bar{B}$。

运算律（和集合论类似）

1. 吸收律：$A \cup \Omega = \Omega, A \cap \Omega = A, A \cup \varnothing = A, A \cap \varnothing = \varnothing, A \cup (AB) = A, A \cap (A + B) = A$；
2. 重余律：$\bar{\bar{A}} = A$；
3. 幂等律：$A \cup A \cup \ldots \cup A = A, A \cap A \cap \ldots \cap A = A$；
4. $A - B = A\bar{B} = A - AB$；
5. 交换律：$A \cap B = B \cap A, A \cup B = B \cup A$；
6. 结合律：$A \cup (B \cup C) = (A \cup B) \cup C, A \cap (B \cap C) = (A \cap B) \cap C$；
7. 分配律：$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), (A \cap B) \cup C = (A \cup C) \cap (B \cup C)$；
8. 对偶律：$\overline{A \cup B} = \bar{A} \cap \bar{B} = \bar{A}\bar{B}, \overline{AB} = \overline{A \cap B} = \bar{A} \cup \bar{B}$；

   对偶律可以推广：$\overline{\cap_{i=1}^{n} A_i} = \cup_{i=1}^{n} \bar{A}_i, \overline{\cup_{i=1}^{n} A_i} = \cap_{i=1}^{n} \bar{A}_i$。

概率

事件 A 发生的可能性大小数值估量称为 A 的概率，记作 $P(A)$。

古典概型

特点：Ω 仅有有限个样本点，每个基本事件发生的可能性大小相同。若满足这两条性质，则为古典概型。

记 $n=\#\Omega, k=\#A$ 分别表示总基本事件个数以及组成 A 的基本事件个数，那么有**古典概型概率公式**：

$$P(A)=\frac{k}{n}=\frac{\#A}{\#\Omega}$$

几何概型

设 Ω 为一个有界区域上的所有点， A 为 Ω上任意子区域，在 Ω 中完全随机任意取点，则取点落入 A 的概率为

$$P(A)=\frac{m(A)}{m(\Omega)}$$

其中，$m(\cdot)$ 为测度；在一维平面上，$m(A)=l_A$为 A 长度；二维平面上，$m(A)=S_A$ 为区域 A 面积。

此时我们称上面这种概型为**几何概型。**

【注意】 $\forall a \in \Omega, P(\{a\})=0$，说明**概率为0的事件不一定为不可能事件**；但**不可能事件概率为0。**

统计概率

设 $A \subseteq \Omega$，在相同条件下重复试验 n 次，其中 A 发生了 n_A 次，则称

$$f_n(A)=\frac{n_A}{n}$$

为 A 在这 n 次试验中发生的**频率**。频率有如下性质：

- $f_n(A) \in [0,1]$；
- $f_n(\Omega)=1, f_n(\varnothing)=0$；
- 可加性：若 $A_1, A_2, \ldots, A_n$两两互斥，则 $f_n(\cup_{i=1}^n A_i)=\sum_{i=1}^n f_n(A_i)$；
- **频率具有稳定性。**

概率的统计定义 在相同条件下重复进行的 n 次试验中，事件 A 发生的频率稳定在一常数 p 附近，在其附近摆动，且随着 n 增大而摆幅越小，称 p 为 A 发生概率，记作 $P(A)=p$。

概率的公理化定义

设 $\mathscr{F}$ 为所有事件的全体，即为**事件域**（集合的集合）。

定义集函数 $P(\cdot): \mathscr{F} \to \mathbb{R}$，将每一个事件 A 映射到一个概率 $P(A)$ 上。

概率的公理化定义 $(\Omega, \mathscr{F})$ 上的概率实质上是 $\mathscr{F}$ 上满足下列三条的集函数：

- **非负性**：$\forall A \in \mathscr{F}, P(A) \geq 0$；
- **归一性**：$P(\Omega)=1$；
- **可列可加性**：若 $A_1, A_2, \ldots, A_n, \ldots$ 为可列的两两互斥事件，则

$$P(\cup_{i=1}^{\infty} A_i)=\sum_{i=1}^{\infty} P(A_i)$$

则称 $(\Omega,\mathscr{F},P)$ 为**概率空间**。显然，$(\Omega,\mathscr{F})$ 上的概率不唯一。

概率的基本性质

- $P(\varnothing)=0$;

  证明：利用可列可加性：由于 $\varnothing\cup\varnothing\cup\ldots\cup\varnothing=\varnothing$，从而 $P(\varnothing)=\sum_{i=1}^{\infty}P(\varnothing)$，因此 $P(\varnothing)=0$。

- **有限可加性**：设 $A_1,A_2,\ldots,A_n$ 为两两互斥事件，则

  $$P(\cup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)$$

  证明：利用可列可加性，令 $A_{n+1}=A_{n+2}=\ldots=\varnothing$，则 $P(\cup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)+\sum_{i=n+1}^{\infty}0$。得证。

  - $0\le P(A)\le 1, P(A)=1-P(\bar{A})$;

    证明：利用有限可加性和归一性，$1=P(\Omega)=P(A)+P(\bar{A})$。

  - $\forall A,B\in\mathscr{F}$，$P(B-A)=P(B)-P(AB)=P(A\cup B)-P(A)$;

    证明：由于 $B=(B-A)+AB, A\cup B=(B-A)+A$，利用有限可加性即可。

  - 特别地，若 $A\subseteq B$，则 $P(B-A)=P(B)-P(A)$;

  - $\forall A,B\in\mathscr{F}$，$P(A\cup B)=P(A)+P(B-A)=P(A)+P(B)-P(AB)\le P(A)+P(B)$；称为**次可加性**或**概率的加法公式**。

  - $\forall A,B,C\in\mathscr{F}$，$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$，称为**容斥原理**或**多除少补原理**。

条件概率

基本定义

一般地，$\forall A,B\in\mathscr{F}$，$P(A)>0$，则称 $P(B|A)$ 为事件 A 发生的条件下，事件 B 发生的**条件概率**，定义为

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

则 $P(\cdot|A)$ 为 $(\Omega,\mathscr{F})$ 上集函数，容易验证满足概率公理化定义的三条性质，因此$P(\cdot|A)$ **是 $(\Omega,\mathscr{F})$ 上的一种概率**，具有上节中叙述的概率的全部性质。

乘法公式

由条件概率定义可得乘法公式：

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

以及推广的乘法公式：

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

全概率公式

一般把 A 设为所关心事件，$B_1,B_2,\ldots,B_n$ 为 n 种可能情况，若满足：

- $\Omega=\cup_{i=1}^{n}B_i$;
- $B_iB_j=\varnothing\ (i,j=1,2,\ldots,n,i\neq j)$

则

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

Bayes 公式

一般把 A 设为所关心事件，$B_1,B_2,\dots,B_n$ 为 n 种可能情况，若满足：

- $\Omega=\cup_{i=1}^{n}B_i$;
- $B_iB_j=\varnothing\ (i,j=1,2,\dots,n,i\neq j)$

则

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

事件的独立性

若 $P(B|A)=P(B)$，则称 A 对 B **没有影响**；如果 A 对 B 没有影响且 B 对 A 没有影响，则 A,B **互不影响。**

设 $A,B\in\mathscr{F}$，若 $P(AB)=P(A)P(B)$，则称事件 A,B 相互**独立。**

独立的性质

- Ω 和 $\varnothing$ 和任何事件都独立；
- 若 A,B 独立，则 $A,\bar{B}$，$\bar{A},B$，$\bar{A},\bar{B}$ 独立。
- 如 $0<P(A)<1$，则 $A,B_{独立}\iff P(B|A)=P(B|\bar{A})=P(B)$。
- 若 $P(A)P(B)>0$ 则 $A,B_{互不相容}\implies A,B_{不相互独立}$

  证明：若 A,B 互不相容，则 $P(AB)=0<P(A)P(B)$，从而不相互独立。

独立性的判定

- 利用定义，$P(AB)=P(A)P(B)$。
- 利用性质3，若 $P(B|A)=P(B)$ 或 $P(B|\bar{A})=P(B)$，则 A,B 相互独立。
- 由题意判断 A,B 是否存在影响。

n **个事件的独立性** 若 n 个事件 $A_1,A_2,\dots,A_n$ 满足如下式子

$$\begin{gathered}P(A_iA_j)=P(A_i)P(A_j)(1\le i<j\le n)\\P(A_iA_jA_k)=P(A_i)P(A_j)P(A_k)(1\le i<j<k\le n)\\\dots\\P(A_1A_2\dots A_n)=P(A_1)P(A_2)\dots P(A_n)\end{gathered}$$

则称 $A_1,A_2,\dots,A_n$ **相互独立。**

n **个事件独立性的判定** 两种方法，① 由定义；② 由题意判定。

n **个事件独立性的性质** 若 $A_1,A_2,\dots,A_n$ 相互独立，则将 n 个事件分成 k 组，对每组事件进行求和、积、差、逆等运算后得到的 k 个事件也相互独立。

2 随机变量及其分布

随机变量

设 E 是随机试验，Ω 为 E 样本空间，$\mathscr{F}$ 为事件域，若有单实值函数 $X(\cdot):\Omega\to\mathbb{R}$，即 $\forall x,\{\omega:X(\omega)\le x\}\in\mathscr{F}$，则称 $X(\omega)$ 为**随机变量。**

分类：离散、非离散（其中一类为连续型）

设 X 为随机变量，则 $\forall x \in \mathbb{R}$，**分布函数**定义为 $F(x) \overset{def}{=} P(\{X \le x\})$.

分布函数的性质

- $0 \le F(x) \le 1, F(+\infty) = 1, F(-\infty) = 0$；
- **单调不减**：$\forall x_1 < x_2, F(x_1) \le F(x_2)$；
- **右连续性**：$F(x+0) = F(x)$。

分布函数的判定 任意函数满足上述三条即可作为分布函数。

分布函数的其他性质 若 X 分布函数 $F(x)$，则

- $P(a \le X \le b) = F(b) - F(a-0)$；
- $P(a < X < b) = F(b-0) - F(a)$；
- $P(X = a) = F(a) - F(a-0)$。

离散型随机变量

基本定义

若随机变量 X 所有可能取值为有限个或可列无穷多个，则称 X 为**离散型随机变量**；用**分布列/分布律**描述分布：

$$P(X = x_i) = p_i \ (i = 1, 2, 3 \ldots, n, \ldots)$$

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

一般令 $x_1, x_2, \ldots, x_n, \ldots$ 递增。

离散型随机变量的分布函数：

$$F(x) = P(X \le x) = \sum_{x_k \le x} P(X = x_k)$$
$$P(X = x_k) = p_k = P(x_{k-1} < X \le x_k) = F(x_k) - F(x_{k-1})$$

常见分布

两点分布（0-1分布）若随机变量 X 满足

$$P(X = k) = p^k (1-p)^{1-k}, (k = 0, 1)$$

则称 X 服从参数为 p 的0-1分布。

二项分布（伯努利分布） 独立重复做 n 次试验，每次有 p 概率成功，$1-p$ 概率失败，最后成功次数 X 服从的分布即为二项分布。即若 X 分布列满足：

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

此时称 X 服从二项分布，记作 $X \sim B(n, p)$。可以发现两点分布即为 $n = 1$ 的二项分布 $B(1, p)$。

二项分布的最可能取值 若 $X \sim B(n, p)$，则若 $(n+1)p$ 为整数，概率在 $k = (n+1)p - 1$ 与 $k = (n+1)p$ 取最大值，否则（$(n+1)p$ 非整数）概率在 $k = [(n+1)p]$ 取最大值，其中 $[\cdot]$ 意为取整。

Poisson分布（泊松分布） 若随机变量 X 分布列为 $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} (k = 0, 1, 2, \ldots)$，则称 X 服从参数为 λ 的泊松分布，记作 $X \sim \pi(\lambda)$ 或 $X \sim P(\lambda)$。

Poisson定理 设 $\lim_{n\to\infty} np_n = \lambda > 0$，则对于固定 k，有

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

当 $n > 100, p < 0.05$ 时可以认为 $C_n^k p_n^k (1-p_n)^{n-k} \approx e^{-\lambda}\frac{\lambda^k}{k!}$。

连续型随机变量

基本定义

设 X 是一随机变量，$X \sim F(x)$，若 $\exists f(x)$，使得 $\forall x \in \mathbb{R}$，$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)du$，则称 X 为**连续型随机变量**，称 $f(x)$ 为概率密度函数。显然连续型随机变量的分布函数 $F(x)$ 绝对连续且唯一，概率密度函数 $f(x)$ 不唯一。

概率密度函数的性质

- $f(x) > 0$；
- $\int_{-\infty}^{+\infty} = F(+\infty) = 1$；
- $f(x)$ 连续点时有 $F'(x) = f(x)$；
- $f(x_0)\Delta x \approx P(x_0 < X \le x_0 + \Delta x)$；
- $P(a < X \le b) = P(a < X < b) = P(a \le X < b) = P(a \le X \le b) = F(b) - F(a)$；
- $\forall k, P(X = k) = 0$。

概率密度函数的判定 满足性质1，2的函数即可作为概率密度函数。

常见分布

区间 (a,b) 上的均匀分布 若 X 的密度函数 $f(x)$ 满足：

$$f(x) = \frac{1}{b-a} \quad (x \in (a,b))$$
$$f(x) = 0 \quad (x \notin (a,b))$$

则称 X 满足区间 (a,b) 上的均匀分布，记作 $X \sim U(a,b)$。其分布函数 $F(x)$ 为：

$$F(x) = \frac{x-a}{b-a} \quad (x \in (a,b))$$
$$F(x) = 0 \quad (x \le a)$$
$$F(x) = 1 \quad (x \ge b)$$

指数分布 若 X 的密度函数 $f(x)$ 为

$$f(x) = \lambda e^{-\lambda x} \quad (x \ge 0)$$
$$f(x) = 0 \quad (x < 0)$$

则称 X 满足参数为 λ 的指数分布，记作 $X \sim E(\lambda)$。其分布函数 $F(x)$ 为

$$F(x) = 1 - e^{-\lambda x} \quad (x \ge 0)$$
$$F(x) = 0 (x < 0)$$

指数分布的性质

- $P(a < X < b) = e^{-\lambda a} - e^{-\lambda b}$；
- **无记忆性**：若 $X \sim E(\lambda)$，则 $P(t < X \le t + \Delta t | X > t) = P(0 < X \le \Delta t) \quad (\Delta t > 0)$。

正态分布 若随机变量 X 的密度函数 $f(x)$ 为

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

则称 X 服从参数为 μ,σ^2 的正态分布，记为 $X\sim N(\mu,\sigma^2)$。其分布函数 $F(x)$ 为超越函数。

标准正态分布 随机变量 $X\sim N(0,1)$，则其分布函数 $\varphi(x)$ 为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

分布函数 $\Phi(x)$ 满足 $\Phi(x)+\Phi(-x)=1$，因此（标准）正态分布具有对称性。标准正态分布函数可以**查表**。

正态分布转化为标准正态分布 若 $X\sim N(\mu,\sigma^2)$，则 $\frac{X-\mu}{\sigma}\sim N(0,1)$。

随机变量函数的分布

离散型随机变量函数的分布

一般方法 设 X 是离散型随机变量，$Y=g(X)$，则求 Y 的分布的一般方法如下：

- $Y=y_1,y_2,\ldots,y_k$ （求出 Y 所有可能取值）；
- $P(Y=y_i)=P(g(X)=y_i)=\sum_{g(x_j)=y_i}p_j$。

连续型随机变量函数的分布

一般方法 设 X 是连续性随机变量，$Y=g(X)$，则求 Y 分布的一般方法如下：

- 先求分布函数 $F_Y(y)=P(Y\le y)=P(g(X)\le y)=\int_{g(x)\le y}f_X(x)dx$；
- 对于 y 求导即得概率密度函数 $f_Y(y)=\frac{d}{dy}F_Y(y)$。

定理 设 X 具有概率密度函数 $f_X(x)$，$g(x)$ 为 $\mathbb{R}$ 上严格单调可导函数，则 $Y=g(X)$ 密度函数为

$$f_Y(y)=|h'(y)|f_x(h(y))\quad(y\in(\alpha,\beta))$$
$$f_Y(y)=0\quad(y\notin(\alpha,\beta))$$

3 多维随机变量及其分布

二维随机变量及其分布

设 E 是随机试验，Ω 为 E 样本空间，$\mathscr{F}$ 为事件域，若函数 $X\times Y:\Omega\to\mathbb{R}^2$，则称 (X,Y) 为**二维随机变量**。

联合分布函数 设 (X,Y) 为二维随机变量，$\forall(x,y)$，$F(x,y)\overset{def}{=}P(X\le x,Y\le y)$ 为二维随机变量 (X,Y) 的联合分布函数。

联合分布函数的性质

- $0\le F(x,y)\le 1, F(+\infty,+\infty)=1, F(\cdot,-\infty)=F(-\infty,\cdot)=0$;
- $F(x,y)$ 对于 x,y 均单调不减；即固定 x（或 y），$F(x,y)$ 关于 y（或 x）单调不减；
- 对 x,y 均右连续，即 $F(x_0,y)=F(x_0,y+0),F(x,y_0)=F(x+0,y_0)$;
- 对任意 $a<b,c<d$，有

$$F(b,d)-F(b,c)-F(a,d)+F(a,c)=P(a<X\le b,c<Y\le d)\ge 0$$

联合分布函数的判定 满足上述四条性质的 $F(x,y)$ 可以称为联合分布函数。

边缘分布函数 设 (X,Y) 联合分布函数为 $F(x,y)$，则 X,Y 各自的分布函数称为边缘分布函数，为

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

联合分布函数唯一确定边缘分布函数，反之不真。

二维离散型随机变量

若二维随机变量的所有可能取值为有限多个或无穷可列多个，则称 (X,Y) 为**二维离散型随机变量。**
(X,Y) 的联合分布律（联合分布列）为：

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

Y/P/X	x_1	x_2	…	x_n
y_1	$p_{1,1}$	$p_{2,1}$	…	$p_{n,1}$
y_2	$p_{1,2}$	$p_{2,2}$	…	$p_{n,2}$
…	…	…	…	…
y_m	$p_{1,m}$	$p_{2,m}$	…	$p_{n,m}$

性质 $p_{i,j}>0,\quad\sum_i\sum_j p_{i,j}=1$。

判定 若满足如上性质可作为二位离散型随机变量的联合分布列。

边缘分布列 $P(Y=y_j)=\sum_i p_{i,j},\quad P(X=x_i)=\sum_i p_{i,j}$。

二维连续型随机变量

基本定义 设 (X,Y) 联合分布函数 $F(x,y)$，若 $\exists f(x,y)\ge 0$，使得
$\forall x,y,F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(x,y)dxdy$，则称 (X,Y) 为二维连续型随机变量，$f(x,y)$ 称为其联合密度函数。

联合密度函数的性质

- $f(x,y)\ge 0$;
- $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(u,v)dudv=F(+\infty,+\infty)=1$;
- 设 G 为平面上区域，则 $P((x,y)\in G)=\iint_G f(u,v)dS$;
- 若 $f(x,y)$ 在 (x_0,y_0) 处二元连续，则 $f(x_0,y_0)=\frac{\partial^2}{\partial x\partial y}F(x,y)|_{x=x_0,y=y_0}=\frac{\partial^2}{\partial y\partial x}F(x,y)|_{x=x_0,y=y_0}$。

联合密度函数的判定 满足性质1，2的函数即可作为联合密度函数。

边缘密度函数 $f_X(x)=F_X'(x)=\int_{-\infty}^{+\infty}f(x,y)dy,\quad f_Y(y)=F_Y'(y)=\int_{-\infty}^{+\infty}f(x,y)dx$。

常见分布

均匀分布 设 D 为平面上一个有界区域，若 $f(x,y)$ 满足

$$f(x,y)=\frac{1}{S_D}((x,y)\in D)$$
$$f(x,y)=0((x,y)\notin D)$$

则称 (X,Y) 服从区域 D 上均匀分布，记作 $(X,Y)\sim U(D)$。

二维正态分布 若 (X,Y) 具有概率密度

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

其中，$\sigma_1>0,\sigma_2>0,\rho\in(-1,1)$。

则称 (X,Y) 服从参数为 $\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho$ 的二维正态分布，记作 $(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$。

则边缘密度函数 $f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$， $f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$ 。

注 若 X，Y 均服从正态分布，那么 (X,Y) 不一定服从二维正态分布！

反例 $f(x,y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}(1+\sin x\sin y)$，容易发现 $X,Y\sim N(0,1)$。

二维随机变量的条件分布

问题 已知 (X,Y) 联合分布，求 $X=x_0$ 下 Y 的分布。

二维离散型随机变量

一般方法 设 $P(X=x_i,Y=y_j)=p_{i,j}$，且 $p_{i,\cdot}=P(X=x_i)=\sum_j p_{i,j}>0$，则

$$P(Y=y_j|X=x_i)=\frac{P(X=x_i,Y=y_j)}{P(X=x_i)}=\frac{p_{i,j}}{p_{i,\cdot}}=\frac{p_{i,j}}{\sum_{j'}p_{i,j'}}$$

二维连续型随机变量

一般方法 设 $(X,Y)\sim f(x,y),X\sim f_X(x),Y\sim f_Y(y)$，则 $X=x$ 下 Y 分布为：

$$F_Y(y|X=x)=\frac{\int_{-\infty}^{y}f(x,t)dt}{f_X(x)}$$
$$f_Y(y|X=x)=\frac{f(x,y)}{f_X(x)}$$

注：$F_{X|Y}(x|y),f_{X|Y}(x|y)$ 是 x 的函数， y 为常数！

性质

- $f_X(x|Y=y_0)=kf(x,y_0)$，则 $k=\frac{1}{f_Y(y_0)}$；
- 连续情况下的条件概率乘法公式 $f(x,y)=f_X(x)f_{Y|X}(y|x)=f_Y(y)f_{X|Y}(x|y)$；
- 连续情况下的全概率公式 $f_X(x)=\int_{-\infty}^{+\infty}f_Y(y)f_X(x|Y=y)dy$；
- 二维正态分布的条件分布仍然是正态分布！

随机变量的独立性

离散型随机变量

对二维离散型随机变量，若满足对于任意 (x_i,y_j) 都有 $P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)$ ，即 $p_{i,j}=p_{i,\cdot}p_{\cdot,j}$，则称 X,Y **互相独立。**

若 X,Y 互相独立，则 $P(X=x_i|Y=y_j)=P(X=x_i)$，即已知 Y 取值不影响 X 的分布。

连续型随机变量

设 (X,Y) 联合密度函数 $f(x,y)$ ，边缘密度函数 $f_X(x),f_Y(y)$，若 $f(x,y)=f_X(x)f_Y(y)$ 在 $f(x,y)$ 的连续点都成立，则称 X,Y **互相独立。**

独立性质及判定：X,Y 独立 $\Longleftrightarrow f_X(x|Y=y)=f_X(x)\Longleftrightarrow f_Y(y|X=x)=f_Y(y)$。

独立必要条件：区域为直矩形区域（根据定义可得）。

判定定理 设 $f(x,y)$ 为 (X,Y) 联合密度函数，则 X,Y 相互独立的充要条件为存在非负可积 $r(x),g(y)$ ，使得 $f(x,y)=r(x)g(y)$ 在 $f(x,y)$ 一切连续点成立。

二维正态分布独立的判定 $\rho=0$。

一般随机变量

一般地，若对任意 x, y 均有 $P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$，即 $F(x,y) = F_X(x)F_Y(y)$ ，则称 X, Y **互相独立**。实际上，该定义等价于任意选取一个矩形区域满足相乘条件（即等价于 $\forall a < b, c < d$，有 $P(a \le X \le b, c \le Y \le d) = P(a \le X \le b)P(c \le Y \le d)$）。

等价描述 实际上，还等价于 $\forall B_x, B_y, P(X \in B_x, Y \in B_y) = P(X \in B_x)P(Y \in B_y)$。

性质 如果 X, Y 独立，则随机变量的函数 $g(X), h(Y)$ 独立。

随机变量相互独立性推广为 n 维 如果 n 维随机变量 $(X_1, X_2, \ldots, X_n)$ 满足

- $P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n) = \prod_{i=1}^n P(X_i = x_i)$
- 或 $F(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n F_{X_i}(x_i)$

则称 $X_1, X_2, \ldots, X_n$ 相互独立。

多维随机变量函数的分布

离散型随机变量

具有可加性的两个分布的叠加

- 若 $X_1 \sim B(n,p), X_2 \sim B(m,p)$，则 $X_1 + X_2 \sim B(n+m, p)$;
- 若 $X_1 \sim P(\lambda_1), X_2 \sim P(\lambda_2)$，则 $X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$。

一般方法 当 (X, Y) 为离散型随机变量时， $Z = g(X, Y)$ 也是离散型随机变量，其中

$$P(Z = z_k) = P(g(X,Y) = z_k) = \sum_{g(x_i, y_j) = z_k} P(X = x_i, Y = y_j) = \sum_{g(x_i, y_j) = z_k} p_{i,j}$$

连续型随机变量

一般方法 当 $(X, Y) \sim f(x, y), Z = g(X, Y)$，则

$$F_Z(z) = P(Z \le z) = P(g(X,Y) \le z) = \iint_{g(X,Y) \le z} f(x,y)dxdy$$

进而，$f_Z(z) = F'_Z(z)$ 即得概率密度函数。

$Z = X + Y$ **的密度函数**

$$F_Z(z) = P(X + Y \le z) = \int_{x+y \le z} f(x,y)dxdy = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} f(x,y)dy$$
$$f_Z(z) = F'_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x)dx = \int_{-\infty}^{+\infty} f(z-y, y)dy$$

当 X, Y 独立时， $f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x) = (f_X * f_Y)(z)$ 为**卷积**。

正态分布的可加性 若 X, Y 独立，且 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$，那么 $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$。

$Z = X/Y$ **的密度函数**

$$F_Z(z) = P(X/Y \le z) = \int_{x/y \le z} f(x,y)dxdy = \int_0^{+\infty} \int_{-\infty}^{yz} f(x,y)dxdy + \int_{-\infty}^0 \int_{yz}^{+\infty} f(x,y)dxdy$$
$$f_Z(z) = F'_Z(z) = \int_0^{+\infty} yf(yz,y)dy - \int_{-\infty}^0 yf(yz,y)dy = \int_{-\infty}^{+\infty} |y|f(yz,y)dy$$

$Z = X^2 + Y^2$ **密度函数**

$$F_Z(z) = P(X^2+Y^2 \le z) = \iint_{x^2+y^2\le z} f(x,y)dxdy = \int_0^{\sqrt{z}} rdr \int_0^{2\pi} f(r\cos\theta, r\sin\theta)d\theta \quad (z \ge 0)$$
$$f_Z(z) = F_Z'(z) = \frac{1}{2}\int_0^{2\pi} f(\sqrt{z}\cos\theta, \sqrt{z}\sin\theta)d\theta \quad (z \ge 0)$$

特别地，n 个独立的标准正态分布的平方和为自由度为 n 的**卡方分布**。

$M = \max(X,Y), m = \min(X,Y)$ **极值分布**

$$F_M(z) = P(\max(X,Y) \le z) = P(X \le z, Y \le z) = F(z,z)$$
$$F_m(z) = P(\min(X,Y) \le z) = 1 - P(X > z, Y > z)$$

注 可以推广至 n 个变量的极值分布。

4 随机变量的数字特征

数学期望

离散型随机变量的数学期望 设 X 为离散型随机变量，其分布律为 $P(X = x_k) = p_k \quad (k = 1,2,\dots)$ ，若无穷级数 $\sum_{k=1}^{+\infty} p_k x_k$ 绝对收敛，则称其和为随机变量 X 的数学期望，记为 $EX = \sum_{k=1}^{+\infty} p_k x_k$。

二项分布的数学期望

$$X \sim B(n,p)$$
$$EX = \sum_{k=0}^{n} kC_n^k p^k (1-p)^{n-k} = \sum_{k=1}^{n} n\frac{(n-1)!}{(n-k)!(k-1)!} p \cdot p^{k-1}(1-p)^{n-k} = np\sum_{k=1}^{n} C_{n-1}^{k-1}p^{k-1}(1-p)^{n-k}$$
$$EX = np\sum_{k=0}^{n-1} C_{n-1}^{k}p^{k}(1-p)^{n-1-k} = np$$

泊松分布的数学期望

$$X \sim P(\lambda)$$
$$EX = \sum_{i=0}^{\infty} i \cdot \frac{\lambda^i}{i!}e^{-\lambda} = e^{-\lambda}\lambda\sum_{i=1}^{\infty}\frac{\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!}} = e^{-\lambda}\lambda e^{\lambda} = \lambda$$

连续型随机变量的数学期望 设连续型随机变量的概率密度为 $f(x)$，若积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛，则称此积分的值为 X 的数学期望，记作 $EX = \int_{-\infty}^{+\infty} xf(x)dx$。

正态分布的数学期望 若 $X \sim N(\mu, \sigma^2)$，则 $EX = \mu$。

均匀分布的数学期望 若 $X \sim U(a,b)$，则 $EX = \frac{a+b}{2}$。

指数分布的数学期望（利用分部积分）

$$X \sim E(\lambda)$$
$$EX = \int_0^{+\infty} xf(x)dx = \int_0^{+\infty} x \cdot \lambda e^{-\lambda x}dx = \dots = \frac{1}{\lambda}$$

随机变量函数的期望 设 Y 为 X 的分段连续函数，若 $Y = g(X)$，$E(Y)$ 存在，则

- 对离散型随机变量 X，若 $P(X = x_k) = p_k$，则 $EY = E(g(X)) = \sum_k g(x_k)p_k$；
- 对连续型随机变量 X，若概率密度为 $f(x)$，则 $EY = E(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$。

设 Z 为 X,Y 的函数，$Z = g(X,Y)$ 且 g 连续，则

- 对离散型随机变量，$EZ = \sum_i\sum_j g(x_i,y_j)p_{i,j}$；
- 对连续型随机变量，$EZ = \iint_{\mathbb{R}^2} g(x,y)f(x,y)dxdy$。

几个重要的随机变量函数期望的定义

- k 阶原点矩：$E(X^k)$；
- k 阶绝对原点矩：$E(|X|^k)$；
- k 阶中心矩：$E((X-EX)^k)$；
- 方差（2 阶中心矩）：$E((X-EX)^2)$；
- 偏度（3 阶中心矩）：$E((X-EX)^3)$；
- 峰度（4 阶中心矩）：$E((X-EX)^4)$。

数学期望的性质 设 X,Y 为随机变量，a,b,c 为常数

- X 期望存在 $\Longleftrightarrow E(|X|)<+\infty$；
- $X\le Y$，则 $E(X)\le E(Y)$；
- $a\le X\le b$，则 $a\le E(X)\le b$；
- **线性性**：$E(aX+bY+c)=aE(x)+bE(y)+c$；

  线性性推广：$E(\sum_{i=1}^n a_iX_i)=\sum_{i=1}^n a_iE(X_i)$；
- 若 X,Y 独立，则 $E(XY)=E(X)E(Y)$，**逆命题不真**；

  推广：若 $X_1,X_2,\dots X_n$ 相互独立，则 $E(\prod_{i=1}^n X_i)=\prod_{i=1}^n E(X_i)$。

方差

设 X 为随机变量，若 EX^2 绝对可积，则称 $E(X-EX)^2$ 为 X 的**方差**，记为 DX，反映了 X 偏离均值程度。

标准差（均方差）定义为：$\sigma_X=\sqrt{DX}$。

计算方法

$$DX=E(X-EX)^2=E(X^2-2EX\cdot X+(EX)^2)=E(X^2)-(EX)^2$$

- 由于 $DX\ge 0$，有 $E(X^2)\ge(EX)^2$。
- 可以通过 EX,DX 计算 EX^2：$EX^2=(EX)^2+DX$。

二项分布的方差

$$X\sim B(n,p)$$
$$EX^2=\sum_{k=0}^n k^2C_n^kp^k(1-p)^{n-k}=\sum_{k=2}^n k(k-1)\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}+\sum_{k=1}^n k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$
$$EX^2=n(n-1)p^2+np$$
$$DX=EX^2-(EX)^2=n(n-1)p^2+np-n^2p^2=np-np^2=np(1-p)$$

泊松分布的方差

$$X\sim P(\lambda)$$
$$DX=EX^2-(EX)^2=\sum_{k=0}^\infty k^2\frac{\lambda^k}{k!}e^{-\lambda}-\lambda^2=e^{-\lambda}\left(\lambda^2\sum_{k=2}^\infty\frac{\lambda^{k-2}}{(k-2)!}+\lambda\sum_{k=1}^\infty\frac{\lambda^{k-1}}{(k-1)!}\right)-\lambda^2$$
$$DX=\lambda^2+\lambda-\lambda^2=\lambda$$

正态分布的方差 若 $X\sim N(\mu,\sigma^2)$，则 $DX=\sigma^2$。

均匀分布的方差 若 $X\sim U(a,b)$，则 $DX=\frac{(b-a)^2}{12}$。

指数分布的方差 （利用分部积分）

$$X \sim E(\lambda)$$
$$DX^2 = \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx - \left(\frac{1}{\lambda}\right)^2 = \ldots = \frac{1}{\lambda^2}$$

方差的性质 设 X, Y 为随机变量，a, b, c 为常数。

- $D(c) = 0$;
- $D(cX) = c^2 D(X)$;
- 若 X, Y 独立，则 $D(X \pm Y) = D(X) + D(Y)$;
- 一般地，$D(X \pm Y) = DX + DY \pm 2E((X - EX)(Y - EY))$（协方差）；
- 若 $X_1, X_2, \ldots, X_n$ 相互独立，则 $D(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i^2 D(X_i)$;
- $D(X) = E(X - EX)^2 \leq E(X - a)^2$

  证明：$g(c) = E(X - c)^2 = E(X^2) - 2cEX + c^2$，当 $c = EX$ 时 $g(c)$ 最小。
- $DX = 0 \Longleftrightarrow \exists c, P(X = c) = 1$（几乎处处是常数）。

标准化随机变量 随机变量 X 的标准化随即变量 X^* 定义为：

$$X^* = \frac{X - EX}{\sqrt{DX}} \quad (DX > 0)$$

那么，$E(X^*) = 0, D(X^*) = 1$。

协方差和相关系数

协方差 若 X, Y 为随机变量，定义 **协方差** 为 $\mathrm{cov}(X, Y) = E(X - EX)(Y - EY)$。特别地，$DX = \mathrm{cov}(X, X)$。称半正定矩阵 $A \in \mathbb{R}^{2\times 2}$ 为 X, Y 的**协方差矩阵**，其中 $a_{1,1} = D(X)$，$a_{2,2} = D(Y)$，$a_{1,2} = a_{2,1} = \mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$。

一般地，n 维随机变量 $(X_1, X_2, \ldots, X_n)$ 存在实对称半正定协方差矩阵 $\Sigma_n = (\mathrm{cov}(X_i, X_j))_{n\times n}$。

相关系数 若 X, Y 为随机变量且 $DX > 0, DY > 0$，则相关系数 $\rho_{X,Y}$ 定义为:

$$\rho_{X,Y} = \frac{\mathrm{cov}(X, Y)}{\sqrt{DX}\sqrt{DY}}$$

- 相关系数 $\rho_{X,Y}$ 为无量纲量;
- $|\rho_{X,Y}|$ 的大小反映了 X, Y 背后的线性关系的强弱。

协方差的计算

- 利用定义，计算 $E(X - EX)(Y - EY)$;
- 用公式计算：$E(X - EX)(Y - EY) = E(XY) - EX \cdot EY$。

协方差的性质

- **对称性**：$\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$;
- $\mathrm{cov}(aX, bY) = ab\mathrm{cov}(X, Y)$;
- $\mathrm{cov}(X + Y, Z) = \mathrm{cov}(X, Z) + \mathrm{cov}(Y, Z)$;
- $\mathrm{cov}(X, X) = DX$;
- $\mathrm{cov}(aX + bY, cX + dY) = acDX + (bc + ad)\mathrm{cov}(X, Y) + bdDY$;
- $D(aX + bY) = \mathrm{cov}(aX + bY, aX + bY) = a^2 DX + b^2 DY + 2ab\mathrm{cov}(X, Y)$。

Cauchy-Schwartz不等式

$$\mathrm{cov}^2(X, Y) \leq DXDY$$

证明：$f(t)=D(X+tY)=(DX)^2+2t\mathrm{cov}(X,Y)+t^2(DY)^2\geq 0$，故 $\Delta=4\mathrm{cov}^2(X,Y)-4DXDY\leq 0$

当 $t=t^*=-\frac{\mathrm{cov}(X,Y)}{DY}$ 时，$D(X+tY)$ 最小值 $D(X+t^*Y)=DX(1-\rho^2_{X,Y})$。

二维正态分布的相关系数 若 $(X,Y)\sim N(\mu_1,\sigma_1;\mu_2,\sigma_2;\rho)$，则 $\rho_{X,Y}=\rho$。

相关系数的性质

- $|\rho_{X,Y}|\leq 1$（由 Cauchy-Schwartz 不等式直接导出）；
- $\rho_{X,Y}=0$ 称 X,Y **不相关**；$\rho_{X,Y}>0$ 称 X,Y **正相关**；$\rho_{X,Y}<0$ 称 X,Y **负相关**；$\rho_{X,Y}=1$ 称 X,Y **完全正相关**；$\rho_{X,Y}=-1$ 称 X,Y **完全负相关**。
- $|\rho_{X,Y}|=1\iff D(X+t^*Y)=DX(1-\rho^2_{X,Y})=0\iff\exists c,P(X+t^*Y=c)=1$，几乎在一条直线上，称**完全线性相关**。
- X,Y 不相关 $\iff\rho_{X,Y}=0\iff\mathrm{cov}(X,Y)=0\iff E(XY)=EXEY\iff D(X\pm Y)=DX+DY$
- **【注】** 独立一定不相关，不相关不一定独立！！
  - 对于二维正态分布 $N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$，$X,Y$ 独立 $\iff X,Y$ 不相关 $\iff\rho=0$

Chebyshev 不等式

一般地，有

$$P(|X-EX|\geq\varepsilon)\leq\frac{DX}{\varepsilon^2}\quad(\varepsilon>0)$$
$$P(|X-EX)\leq\varepsilon)\geq 1-\frac{DX}{\varepsilon^2}\quad(\varepsilon>0)$$

证明：

$$P=P(|X-EX|\geq\varepsilon)$$
$$DX=\left(\int_{-\infty}^{EX-\varepsilon}+\int_{EX-\varepsilon}^{EX+\varepsilon}+\int_{EX+\varepsilon}^{+\infty}\right)(X-EX)^2f(x)dx\geq\left(\int_{-\infty}^{EX-\varepsilon}+\int_{EX+\varepsilon}^{+\infty}\right)\varepsilon^2f(x)dx=\varepsilon^2P$$

5 大数定律与中心极限定理

大数定律

Bernoulli大数定理 设 n_A 为 n 次独立重复试验中事件 A 发生次数，p 是每次试验中 A 发生的概率，则 $\forall\varepsilon>0$，

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

或

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

依概率收敛 设 $Y_1,Y_2,\dots,Y_n$ 为一随机变量序列，a 是一常数，且 $\forall\varepsilon>0$ 有

$$\lim_{n\to\infty}P\left(|Y_n-a|\geq\varepsilon\right)=0$$

或

$$\lim_{n\to\infty}P\left(|Y_n-a|<\varepsilon\right)=1$$

则称随机变量序列 $Y_1, Y_2, \ldots, Y_n$ 依概率收敛至 a，记作 $Y_n \stackrel{p;\, n\to\infty}{\longrightarrow} a$。

服从大数定律 若随机变量序列 $X_1, X_2, \ldots, X_n$ 具有如下性质则称这个随机变量序列服从大数定律：

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

即 $\frac{1}{n}\sum_{k=1}^{n} X_k$ 依概率收敛到 $\frac{1}{n}\sum_{k=1}^{n} E(X_k)$。

若 $X_1, X_2, \ldots, X_n$ 独立同分布，则上式等价于

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

即 $\frac{1}{n}\sum_{k=1}^{n} X_k$ 依概率收敛到 EX。

Chebyshev 大数定律

设随机变量序列 $X_1, X_2, \ldots, X_n$ 两两不相关，它们的方差存在，而且有共同上界，即

$$E(X_k) = \mu_k, \quad D(X_k) = \sigma_k^2 \le \sigma^2 < +\infty \quad (k = 1, 2, \ldots, n)$$

则这个随机变量序列服从大数定律，即

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

证明：令 $Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$，则 $E(Y_n) = \frac{1}{n}\sum_{k=1}^{n} \mu_k$。

$$D(Y_n) = \frac{1}{n^2} D\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) \le \frac{\sigma^2}{n^2}\cdot n = \frac{\sigma^2}{n}$$

根据 Chebyshev 不等式，$P(|Y_n - E(Y_n)| < \varepsilon) \ge 1 - \frac{DY_n}{\varepsilon^2} = 1 - \frac{\sigma^2}{n\varepsilon^2} \to 1\ (n \to \infty)$。

即满足大数定律。

注意：所有条件可以削弱至 **Markov 条件**

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

Khintchine 大数定律 设 $X_1, X_2, \ldots, X_n$ 独立同分布，若它们期望存在，$E(X_i) = \mu$，则该序列服从大数定律，即 $\frac{1}{n}\sum_{k=1}^{n} X_k$ 依概率收敛到 μ，即

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

中心极限定理

Lindeberg-L é vy 中心极限定理 设 $X_1, X_2, \ldots, X_n$ 独立同分布，且

$$E(X_i) = \mu, \quad D(X_i) = \sigma^2 < +\infty$$

则 $\forall x$，

$$\lim_{n\to\infty} P\left(\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x\right) = \Phi(x)$$

即 $\sum_{i=1}^n X_i$ 近似服从 $N(n\mu, n\sigma^2)$，$Y_n = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma}$ 在 n 很大时近似服从于 $N(0,1)$。

6 数理统计的基本概念

总体与样本

总体 研究对象的全体（研究对象某项数量指标的全体）组成的集合。

个体 组成总体的每一个个体。

总体分布 总体中每个个体数量指标不同，背后有一定的分布，称为总体分布。（把总体看成一个随机变量 X，总体分布即是随机变量 X 的分布）。

样本 从总体中随机抽取部分个体组成的集合称为样本 $(X_1, X_2, \ldots, X_n)$，n 称为**样本容量**。

样本值（样本观察值）$(x_1, x_2, \ldots, x_n)$ 称为总体 X 的一个容量为 n 的样本观察值，即样本的一组可能取值。

样本空间 样本 $(X_1, X_2, \ldots, X_n)$ 的所有可能取值的集合，记作 $\mathscr{X}$。

简单随机样本（如不加特别说明，样本默认为简单随机样本）设 $(X_1, X_2, \ldots, X_n)$ 为来自 X 的一个样本，若其满足

- $X_1, X_2, \ldots, X_n$ 与 X 同分布；
- $X_1, X_2, \ldots, X_n$ 相互独立；

则称这个样本是简单随机样本。

- 简单随机样本 $(X_1, X_2, \ldots, X_n)$ 的联合分布函数 $F(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n F_X(x_i)$；
- 简单随机样本 $(X_1, X_2, \ldots, X_n)$ 的联合密度函数 $f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n f_X(x_i)$；

统计量 若 $g(x_1, x_2, \ldots, x_n)$ 是不含未知参数的实值连续函数，则随机变量 $g(X_1, X_2, \ldots, X_n)$ 为统计量，$(x_1, x_2, \ldots, x_n)$ 为样本 $(X_1, X_2, \ldots, X_n)$ 观察值，则 $g(x_1, x_2, \ldots, x_n)$ 为统计量 $g(X_1, X_2, \ldots, X_n)$ 观察值。

常用的一些统计量

- 样本均值 $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$，依概率收敛至 EX；
- 样本方差 $s^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$，依概率收敛至 DX；
- 样本均方差 $s = \sqrt{s^2}$，依概率收敛至 $\sqrt{DX}$；
- 样本 k 阶矩 $M_k = \frac{1}{n}\sum_{i=1}^n X_i^k$，依概率收敛至 $E(X^k)$；
- 样本 k 阶中心矩 $CM_k = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^k$，依概率收敛至 $E((X-EX)^k)$。
- 顺序统计量：将 $(X_1, X_2, \ldots, X_n)$ 从小到大排序后得到 $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$，即为**顺序统计量**，$X_{(1)} = \min_{i=1}^n X_i, X_{(n)} = \max_{i=1}^n X_i$，**极差** $D = X_{(n)} - X_{(1)}$。

抽样分布

正态分布

$$X \sim N(\mu, \sigma^2), \quad (X_1, X_2, \ldots, X_n), \quad X_i \sim N(\mu, \sigma^2)$$
$$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i \sim N(\mu, \frac{\sigma^2}{n})$$
$$\sum_{i=1}^n a_i X_i \sim N\left(\mu \sum_{i=1}^n a_i, \sigma^2 \sum_{i=1}^n a_i^2\right)$$

卡方分布

设 $(X_1, X_2, ..., X_n)$ 独立同分布，且 $X_i \sim N(0,1)$，则
$\chi^2 = X_1^2 + X_2^2 + ... + X_n^2 = \sum_{i=1}^n X_i^2 \sim \chi^2(n)$，称 χ^2 服从自由度为 n 的**卡方分布。**

图像（均在 y 轴右半边）

- $n = 1$ 时候递减，类似反比例函数；
- $n = 2$ 时候递减， $\chi^2(2) = E(\frac{1}{2})$；
- $n \geq 3$ 时出现峰，且 n 越大峰越靠后。

性质

- **可加性**：若 $\chi_1^2 \sim \chi^2(n_1), \chi_2^2 \sim \chi^2(n_2)$ 且 χ_1^2, χ_2^2 独立，则 $\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$。
- 若 $X \sim \chi^2(n)$，则（利用了自由度为2的卡方分布是参数为 $\frac{1}{2}$ 的指数分布）

$$EX = E(X_1^2 + X_2^2 + ... + X_n^2) = \sum_{i=1}^n E(X_i^2) = n(EX_i + DX_i) = n(0+1) = n$$

$$DX = D(X_1^2 + X_2^2 + ... + X_n^2) = nD(X_1^2) = \frac{n}{2} D(X_1^2 + X_2^2) = \frac{n}{2} \times \frac{1}{\left(\frac{1}{2}\right)^2} = 2n$$

  同时，这个结果给出了对于服从标准正态分布 $N(0,1)$ 的随机变量 X 的平方的期望为1，方差为2。

查表：若 $P(X > k) = \alpha$，则记 $k = \chi_\alpha^2(n)$，称其为 $\alpha-分位数$。

t-分布

设 $X \sim N(0,1), Y \sim \chi^2(n)$，$X, Y$ 独立，则

$$T = \frac{X}{\sqrt{\frac{Y}{n}}}$$

所服从的分布称为自由度为 n 的 t-分布。记为 $T \sim t(n)$。

图像&性质

- $f(t)$ 为偶函数，图像关于 y 轴对称；
- $n \to +\infty$，$T \overset{近似}{\sim} N(0,1)$。（根据**Khintchine大数定律**，$Y/n$ 依概率收敛到 $E(Y_1^2) = 1$。

查表：若 $P(T > k) = \alpha$，则记 $k = t_\alpha(n)$，称其为 $\alpha-分位数$。

F-分布

设 $X \sim \chi^2(m), Y \sim \chi^2(n)$，且 X, Y 独立，则

$$F = \frac{\frac{X}{m}}{\frac{Y}{n}}$$

服从的分布称为第一自由度 m，第二自由度 n 的 F-分布，记为 $F \sim F(m,n)$。

图像 和卡方分布类似。

性质

- $F \sim F(m,n)$，则 $\frac{1}{F} \sim F(n,m)$；
- $T \sim t(n)$，则 $T^2 \sim F(1,n)$。

查表：若 $P(F > k) = \alpha$，则记 $k = F_\alpha(n)$，称其为 $\alpha-分位数$。

分位数性质

- $F_{1-\alpha}(m,n)=\frac{1}{F_\alpha(n,m)}$;

  证明：$X\sim F(m,n)$，令 $F_{1-\alpha}(m,n)=k$，则 $P(X>k)=1-\alpha$，从而 $P(X\le k)=\alpha$，从而 $P\left(\frac{1}{X}\ge\frac{1}{k}\right)=\alpha$，而 $\frac{1}{X}\sim F(n,m)$，从而 $F_\alpha(n,m)=\frac{1}{k}$，从而$F_{1-\alpha}(m,n)=\frac{1}{F_\alpha(n,m)}$。

- $[t_{1-\frac{\alpha}{2}}(n)]^2=F_\alpha(1,n)$;

  证明：$X\sim t(n)$，设 $k=t_{1-\frac{\alpha}{2}}(n)$，则 $P(X>k)=1-\frac{\alpha}{2}$。容易知道 $k\le 0$（因为 $P(X>k)\ge\frac{1}{2}$，t-分布具有对称性），从而 $P(X\le k)=\frac{\alpha}{2}$，$P(|X|>-k)=\alpha$，$P(X^2>k^2)=\alpha$。又 $X^2\sim F(1,n)$，从而 $k^2=F_\alpha(1,n)$，于是 $[t_{1-\frac{\alpha}{2}}(n)]^2=F_\alpha(1,n)$。

正态总体的抽样分布

一个正态总体的抽样分布

$$X\sim N(\mu,\sigma^2),\quad EX=\mu,\quad DX=\sigma^2,\quad (X_1,X_2,\dots,X_n)$$
$$\bar{X}=\frac{1}{n}\sum_{i=1}^n X_i\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$
$$\frac{1}{\sigma^2}\sum_{i=1}^n(X_i-\mu)^2=\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$$
$$\frac{n-1}{\sigma^2}s^2=\frac{1}{\sigma^2}\sum_{i=1}^n(X_i-\bar{X})^2\sim\chi^2(n-1)$$
$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$$
$$\frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}}=\frac{\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{1}{n-1}\cdot\frac{n-1}{\sigma^2}s^2}}\sim t(n-1)$$

重要性质：$\bar{X}$ 和 s 相互独立。

两个独立正态总体的分布情况（X,Y 独立）

$$X\sim N(\mu_1,\sigma_1^2),\quad Y\sim N(\mu_2,\sigma_2^2),\quad (X_1,X_2,\dots,X_n),\quad (Y_1,Y_2,\dots,Y_m)$$
$$\bar{X}-\bar{Y}\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}\right)$$

若 $\sigma_1=\sigma_2=\sigma$，则

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}\sigma}\sim N(0,1)$$
$$s_w=\sqrt{\frac{n-1}{n+m-2}s_1^2+\frac{m-1}{n+m-2}s_2^2}$$
$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}s_w}=\frac{\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}\sigma}}{\sqrt{\frac{1}{n+m-2}\left(\frac{n-1}{\sigma_1^2}s_1^2+\frac{m-1}{\sigma_2^2}s_2^2\right)}}\sim t(n+m-2)$$

另外还有

$$\frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} = \frac{\frac{1}{n-1}\frac{n-1}{\sigma_1^2}s_1^2}{\frac{1}{m-1}\frac{m-1}{\sigma_2^2}s_2^2} \sim F(n-1,m-1)$$

$$\frac{s_1^2}{s_2^2} \sim F(n-1,m-1) \quad (\sigma_1^2=\sigma_2^2)$$

7 参数估计

点估计法

设总体 $X \sim F(x;\theta_1,\theta_2,\dots,\theta_k)$，其中 $\theta_1,\theta_2,\dots,\theta_k$ 为未知参数。用 k 个统计量估计 $\theta_1,\theta_2,\dots,\theta_k$ 的方法称为**点估计法。**

矩法

$$\frac{1}{n}\sum_{i=1}^n X_i = \bar{X} \approx EX = \mu_1(\theta_1,\theta_2,\dots,\theta_k)$$
$$\frac{1}{n}\sum_{i=1}^n X_i^2 = M_2 \approx EX^2 = \mu_2(\theta_1,\theta_2,\dots,\theta_k)$$
$$\cdots$$
$$\frac{1}{n}\sum_{i=1}^n X_i^k = M_k \approx EX^k = \mu_k(\theta_1,\theta_2,\dots,\theta_k)$$

一般地，

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^n X_i = \bar{X}$$
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2 = CM_2$$

注：如果上述 k 个方程中有无用方程（恒等），则继续往下列写，直到有 7

极大似然估计

离散型下的似然函数 一般地，设 X 为离散型总体，其分布律为 $P(X=x)=p(x;\theta)\ (x=\mu_1,\mu_2,\dots,\mu_m,\dots)$，$(X_1,X_2,\dots,X_n)$ 为总体样本，似然函数如下定义：

$$P(X=x_1,X=x_2,\dots,X=x_n) = \prod_{i=1}^n P(X=x_i) = \prod_{i=1}^n p(x_i;\theta) \overset{def}{=} L(x_1,x_2,\dots,x_n;\theta)$$

离散型下的对数似然函数 一般地，对数似然函数如下定义：

$$\ln L(x_1,x_2,\dots,x_n;\theta) = \sum_{i=1}^n \ln p(x_i;\theta)$$

连续型下的似然函数 类似地，似然函数如下定义：

$$L(x_1,x_2,\dots,x_n;\theta_1,\theta_2,\dots,\theta_k) \overset{def}{=} \prod_{i=1}^n f(x_i;\theta_1,\theta_2,\dots,\theta_k)$$

连续型下的对数似然函数 一般地，对数似然函数如下定义：

$$\ln L(x_1,x_2,\dots,x_n;\theta_1,\theta_2,\dots,\theta_k) = \sum_{i=1}^n \ln f(x_i;\theta_1,\theta_2,\dots,\theta_k)$$

方法 寻找 $\theta=\hat{\theta}$，使得似然函数（或对数似然函数）$L(\theta)$ （或 $\ln L(\theta)$）取极大值，即

$$L(x_1,x_2,\ldots,x_n;\hat{\theta})=\max_{\theta\in\Theta}\left\{\prod_{i=1}^{n}p(x_i;\theta)\right\}$$

则称这样得到的 $\hat{\theta}=g(x_1,x_2,\ldots,x_n)$ 即为参数的估计值，$g(X_1,X_2,\ldots,X_n)$ 为**极大似然估计量。**

若有多个参数，则对于每个参数求偏导后令其等于0得到一个方程，联立解出估计值即可。

极大似然估计的不变性原理 设 $\hat{\theta}$ 为 θ 的极大似然估计量，$u(\theta)$ 为 θ 的连续函数，则 $u(\hat{\theta})$ 也为 $u(\theta)$ 的极大似然估计量。

泊松分布总体的极大似然估计 设 $X\sim P(\lambda)$，样本 $(X_1,X_2,\ldots,X_n)$，则经过极大似然估计后，有

$$\lambda=\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

正态总体的极大似然估计 设 $X\sim N(\mu,\sigma^2)$，样本 $(X_1,X_2,\ldots,X_n)$，则经过极大似然估计后，有

$$\hat{\mu}=\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

估计量的评判标准

无偏性

设总体 $X\sim F(x;\theta)$，θ 为位置参数， $\hat{\theta}=\hat{\theta}(X_1,X_2,\ldots X_n)$ 为 θ 的估计量（是一个统计量！），若对任意设定的 θ 都有 $E(\hat{\theta})=\theta$，则说 $\hat{\theta}$ 为 θ 的**无偏估计量。**

同一个参数的无偏估计量不唯一

- $\bar{X}$ 为 EX 无偏估计，$\frac{1}{n}\sum_{i=1}^{n}X_i^k$ 为 EX^k 无偏估计；
- $CM_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 不是正态总体 $N(\mu,\sigma^2)$ 中 σ^2 的无偏估计，$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 是 σ^2 的无偏估计。

有效性

设 $\hat{\theta}_1,\hat{\theta}_2$ 都是 θ 的无偏估计量（是统计量！），若 $D(\hat{\theta}_1)<D(\hat{\theta}_2)$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效。

Rao-Cramer不等式 若 $\hat{\theta}$ 为 θ 的无偏估计量，则

$$D(\hat{\theta})\geq\frac{1}{nE\left[\frac{\partial}{\partial\theta}\ln p(X,\theta)\right]^2}\overset{def}{=}D_0$$

有效估计量 能达到 Rao-Crammer 不等式下界的估计量。

一致性

设 $\hat{\theta_n}=\hat{\theta}(X_1,X_2,\ldots,X_n)$ 为 θ 估计量，若对任意 θ，当 $n\to\infty$ 时，θ_n 依概率收敛至 θ，即

$$\lim_{n\to\infty}P(|\hat{\theta_n}-\theta|\geq\varepsilon)=0\quad(\varepsilon>0)$$

则称 $\hat{\theta_n}$ 为 θ 的**一致估计量。**

- M_k 为 EX^k 的一致估计量。
- 矩估计量一般是一致估计量。

- 若 $\hat{\theta}$ 是 θ 的一致估计量，u 为 θ 的连续函数，则 $u(\hat{\theta})$ 为 $u(\theta)$ 的一致估计量。
- $s^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$ 是总体方差的一致估计量。

区间估计

设 $X\sim F(x,\theta)$，θ 未知参数，给定**置信度** 为 $1-\alpha\ (0<\alpha<1)$，若 $\exists\hat{\theta_1},\hat{\theta_2}$，使得 $P(\hat{\theta_1}\le\theta\le\hat{\theta_2})=1-\alpha$，则称区间 $[\hat{\theta_1},\hat{\theta_2}]$ 为参数 θ 的置信度为 $1-\alpha$ 的**置信区间。**

- 置信区间长度 $\hat{\theta_2}-\hat{\theta_1}$ 反映了估计精度，长度越小精度越高；
- 置信度 $1-\alpha$ 反映了可靠度，α 越小精度越高；
- 先保证可靠性再提高精度，通常 $1-\alpha=0.95$；
- 扩大样本容量可以提高精度。

一般方法

- 构造**枢轴量**（仅含有待估参数 θ，分布确定且不依赖于待估参数） $g(X_1,X_2,\dots,X_n;\theta)$；
- 给定置信度 $1-\alpha$，定出常数 a,b 使得 $P(a<g(X_1,X_2,\dots,X_n;\theta)<b)=1-\alpha$ （a,b不唯一）；
- 反解不等式 $a<g(X_1,X_2,\dots,X_n;\theta)<b$ 得到 $\hat{\theta_1}<\theta<\hat{\theta_2}$。

一个正态总体的区间估计

- σ^2 已知，求 μ 区间估计，枢轴量

$$U=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$$

- σ^2 未知，求 μ 区间估计，枢轴量

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

- μ 已知，求 σ^2 区间估计，枢轴量

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

- μ 未知，求 σ^2 区间估计，枢轴量

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

单侧置信区间 对于给定 α，θ 待估，若存在 $\underline{\theta}$ 使得 $P(\underline{\theta}\le\theta)=1-\alpha$，则称 $(\underline{\theta},+\infty)$ 为 θ 的**单侧置信区间**，$\underline{\theta}$ 为**置信下限**；同理有**置信上限**和另一个**单侧置信区间。**

两个正态总体的参数估计（X,Y 独立）

$$X\sim N(\mu_1,\sigma_1^2),\quad Y\sim N(\mu_2,\sigma_2^2),\quad (X_1,X_2,\dots,X_n),\quad (Y_1,Y_2,\dots,Y_m)$$

- σ_1,σ_2 已知，求 $(\mu_1-\mu_2)$ 置信区间，枢轴量

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

- σ_1,σ_2 未知，但有 $\sigma_1^2=\sigma_2^2$ ，求 $(\mu_1-\mu_2)$ 置信区间，枢轴量

$$s_w=\sqrt{\frac{n-1}{n+m-2}s_1^2+\frac{m-1}{n+m-2}s_2^2}$$
$$T=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{s_w\sqrt{\frac{1}{n}+\frac{1}{m}}}\sim t(n+m-2)$$

- σ_1, σ_2 未知，大样本，可以近似估计 $\hat{\sigma_1}^2 = s_1^2, \hat{\sigma_2}^2 = s_2^2$ 后利用 $U \sim N(0,1)$。
- μ_1, μ_2 不要求，求 $\frac{\sigma_1^2}{\sigma_2^2}$ 置信区间，枢轴量

$$F = \frac{\frac{s_1^2}{s_2^2}}{\frac{\sigma_1^2}{\sigma_2^2}} = \frac{\frac{1}{n-1}\frac{n-1}{\sigma_1^2}s_1^2}{\frac{1}{m-1}\frac{m-1}{\sigma_2^2}s_2^2} \sim F(n-1, m-1)$$

8 假设检验

对一个或多个总体的分布或参数作假设。从总体中抽样，用统计的方法对假设合理性做出判断，这类方法称作**假设检验**（统计假设检验）。假设检验分为参数检验和非参数检验。其理论依据为：实际推断原理（小概率原理），即小概率事件在一次试验中不会发生。

一般方法

- 建立假设，H_0 为原假设（偶然因素导致的为原假设，一般为题目所问），H_1 为备选假设（H_0, H_1 不能交换且 H_0, H_1 互斥但并集为全集）；如果 H_0 涉及到复合假设如 $H_0: \mu \geq 70$，应该简单化得到 $H_0': \mu = 70$。
- 选择适当统计量 $T = g(X_1, X_2, \ldots, X_n)$（要求 H_0 成立条件下，T 的分布完全已知）；给定显著性水平 α，依据 $P_{H_0}(T \in W) \leq \alpha$ 构造拒绝域 W。
- 代入数据检验：$\hat{T} = g(X_1, X_2, \ldots, X_n)$，如果 $\hat{T} \in W$，则拒绝 H_0；否则接受 H_0。

注意

- 拒绝域由备选假设决定，α 越大越容易拒绝；
- 两类错误：
  - 第一类错误：拒绝 H_0 但是 H_0 为真；
  - 第二类错误：接受 H_0 但是 H_0 为假；
  - 犯第一类错误概率减小必然导致犯第二类错误概率增大，反之亦然；
  - 应该尽可能避免犯第一类错误，控制犯第一类错误的概率为 α；
  - 要使犯两类错误概率都减小——扩大样本容量。

一个正态总体的参数检验

- 关于 μ 检验：$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$：
  - 方差 σ^2 已知，则

$$U = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \overset{H_0}{\sim} N(0,1)$$

    称为 U-检验法；
  - 方差 σ^2 未知，则

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \overset{H_0}{\sim} t(n-1)$$

    称为 T-检验法；
- 关于 σ^2 的检验：$H_0: \sigma^2 = \sigma_0^2, H_1: \sigma^2 \neq \sigma_0^2$：
  - 均值 μ 已知，则

$$K = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 \overset{H_0}{\sim} \chi^2(n)$$

  - 均值 μ 未知，则

$$K=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\bar{X})^2\overset{H_0}{\sim}\chi^2(n-1)$$

上述均称为 K-检验法。

两个正态总体的参数检验（X,Y 独立）

$$X\sim N(\mu_1,\sigma_1^2),\quad Y\sim N(\mu_2,\sigma_2^2),\quad (X_1,X_2,\dots,X_n),\quad (Y_1,Y_2,\dots,Y_m)$$

- 检验 $(\mu_1-\mu_2)$，$H_0:\mu_1-\mu_2=\delta,H_1:\mu_1-\mu_2\neq\delta$。
  - σ_1,σ_2 已知，则
  $$U=\frac{\bar{X}-\bar{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}\overset{H_0}{\sim}N(0,1)$$
  - σ_1,σ_2 未知，但有 $\sigma_1^2=\sigma_2^2$，则
  $$s_w=\sqrt{\frac{n-1}{n+m-2}s_1^2+\frac{m-1}{n+m-2}s_2^2}$$
  $$T=\frac{\bar{X}-\bar{Y}-\delta}{s_w\sqrt{\frac{1}{n}+\frac{1}{m}}}\overset{H_0}{\sim}t(n+m-2)$$
- 检验 $\frac{\sigma_1^2}{\sigma_2^2}$，$H_0:\frac{\sigma_1^2}{\sigma_2^2}=\delta,H_1:\frac{\sigma_1^2}{\sigma_2^2}\neq\delta$，则
  $$F=\frac{1}{\delta}\frac{s_1^2}{s_2^2}=\sim F(n-1,m-1)$$

  称为 F-检验法。