概率論部分

§1　随機事件及其概率

§1.1　随機事件及其運算

两个事件 A, B 的

- 包含关系 $A \subset B$，表示 A 发生将导致 B 发生，$P(B-A)=P(B)-P(A), P(B) \geqslant P(A)$.
- 和事件 $A \cup B$，表示 A, B 至少有一者发生，$A \cup B = A + \overline{A}B$;
- 积事件 $A \cap B$，表示 A, B 同时发生;
- 差事件 $A - B$，表示 A 发生，B 不发生.
- 互不相容 / **互斥**，$A \cap B = \varnothing$，此时 $P(A \cup B) = P(A+B) = P(A) + P(B)$，是为概率的 **有限可加性**;
- 对立 / 互逆，$A \cap B = \varnothing$ 且 $A \cup B = S$，其中 S 是样本空间，此时，$B = \overline{A}$;
- 独立，$P(AB) = P(A)P(B)$. 多个事件 $A_1, A_2, \ldots, A_n$ 相互独立：若对于任意 k 个其中事件$P(A_{i_1}A_{i_2}\ldots A_{i_k}) = P(A_{i_1})P(A_{i_2})\ldots P(A_{i_k})$. 相互独立则必定两两独立，但是反之不成立 †.

  † : 反例：考虑一个四面体，一面红色，一面黄色，一面蓝色，一面三种颜色都有. 现在掷此四面体. A：朝下面包含红色，B：朝下面包含黄色，C：朝下面包含蓝色. 这三个事件两两独立，但不相互独立.

这些定义和运算规则与集合类似，如性质

$$\text{(De-Morgan)} \qquad \overline{\bigcup_{j=1}^{n} A_j} = \bigcap_{j=1}^{n} \overline{A_j}, \quad \overline{\bigcap_{j=1}^{n} A_j} = \bigcup_{j=1}^{n} \overline{A_j}$$

$$\text{(Jordan)} \qquad P(A \cup B) = P(A) + P(B) - P(AB)$$

$$P\Big(\bigcup_{j=1}^{n} A_j\Big) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leqslant j_1 < j_2 < \cdots < j_k \leqslant n} P(A_{j_1}A_{j_2}\ldots A_{j_k})$$

§1.2 等可能模型

古典概型 试验的样本空间 Ω 有限，$A \subseteq \Omega$，若 Ω 中的每个样本点的发生的可能性相同，则称 $P(A) = \frac{|A|}{|\Omega|}$ 为事件 A 发生的概率．这将问题转化为了求集合中的元素个数，利用排列组合的知识即可解决．古典概型满足有限可加性．

几何概型 对于 $A \subseteq \mathbb{R}^r$，设 A 的体积存在，记为 $m(A)$．试验的样本空间 Ω，$m(\Omega) > 0$，$A \subseteq \Omega$，若每个样本点等可能地落在 Ω 中，则称 $P(A) = \frac{m(A)}{m(\Omega)}$ 为事件 A 发生的概率．一个经典的应用是等待—相遇问题．

概型的差别的根源是等可能假设之间的差别，如 Bertrand 悖论†，不同的等可能假设导致了截然不同的结果，这些结果之间并无对错之分，只是从不同角度看同一个问题得到了不同的结果．

† : 在半径为 1 的圆内任意取一条弦，求弦长度大于等于 $\sqrt{3}$ 的概率．

§1.3 條件概率

条件概率 在条件 A 下 B 发生的概率

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

全概率公式 事件 $A_1, A_2, \ldots, A_n$ 互斥，$B \subseteq \bigcup_{j=1}^n A_j$，有

$$P(B) = \sum_{j=1}^{n} P(A_j) P(B \mid A_j)$$

Bayes 公式 事件 $A_1, A_2, \ldots, A_n$ 互斥，$B \subseteq \bigcup_{j=1}^n A_j$，若 $P(B) > 0$，

$$P(A_j \mid B) = \frac{P(A_j) P(B \mid A_j)}{\sum_{j=1}^{n} P(A_j) P(B \mid A_j)}$$

§2 随機變量與分布

每个随机试验都可以对应一个随机变量，随机试验的结果（样本点）对应其取值.

随机变量与一般的函数有所区别．对于一般函数， 每一个自变量就对应一个函数值；对于随机变量，虽然也如此，但自变量（样本点）能否被取到会受到概率的影响，因此函数值（随机变量的取值）也有其概率.

§2.1 離散型随機變量

随机变量 X 只能取到 **有限** 个或 **可列** 个不同值.

§2.1.1 離散型随機變量的分布

概率分布 分布列(律) $P(X=x_k)=p_k$，满足 $p_k \geqslant 0$，$\sum p_k = 1$.

分布函数 $F(x)=\sum\limits_{j:x_j\leqslant x} p_j$，这是单调递增的阶梯函数，且函数值的跳跃发生在所有 x_k 处，跳跃的幅度为 p_k.

§2.1.2 離散型随機變量的函數的分布

已知离散型随机变量 X 的概率分布，$Y=g(X)$，用 **合并分布列** 求 Y 分布．具体地，假设 X 的分布列 $P(X=x_k)=p_k$，合并 $g(x_1), g(x_2), \ldots$ 中的相同值作为 Y 的不同取值 $y_1, y_2, \ldots$，则 Y 的分布列

$$P(Y=y_j)=\sum_{i:g(x_i)=y_j} P(X=x_i)$$

§2.1.3 常見離散概率分布

1. (0-1) 分布（兩點分布）

- **模型** X 的值为一个随机事件中发生的事件数，这个事件发生的概率为 p，且 X 只能取 0 或 1，则 $X \sim b(1, p)$.
- **分布** $P(X=1)=p$, $P(X=0)=1-p$
- **期望和方差** $\mathrm{E}X=p$，$\mathrm{D}X=p(1-p)$

2. 二項分布

- **模型**（有放回摸球）X 为 n 次 **独立重复** 随机事件（n 重伯努利试验）中发生的事件数，这个事件每次发生的概率都是 p. 则 $X \sim b(n, p)$.
- **分布** $\boxed{P(X=k)=\binom{n}{k} p^k(1-p)^{n-k}}, \quad k=0,1,\ldots,n$
- **期望和方差** $\mathrm{E}X=np$，$\mathrm{D}X=np(1-p)$

$\binom{n}{k}$ 是二项式系数，即 C_n^k.

独立条件下二项分布对于其中的参数 n 具有可加性：若 $X_1 \sim b(n_1, p), X_2 \sim b(n_2, p)$，且 X_1, X_2 独立，则有 $X_1+X_2 \sim b(n_1+n_2, p)$. 服从二项分布的随机变量可以看作是 n 个服从 (0-1) 分布的独立的随机变量之和.

二项分布的图形特征：对于固定 n 及 p，当 k 增加时，概率 $P(X=k)$ 先随之增加直至达到最大值，随后单调减少.

*3. 多项分布

- **模型**（有放回摸球）$A_1, A_2, \dots, A_r$ 为完备事件组（并是 Ω，且两两互斥）. 独立重复试验 n 次，X_i 表示事件 A_i 的发生次数，$P(A_i) = p_i$.
- **分布** $P(X_1 = k_1, X_2 = k_2, \dots, X_r = k_r) = \binom{n}{k_1\ k_2\ \dots\ k_r}^{\dagger} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}, \quad k_i \geqslant 0, \sum k_i = n$
- **例** 兹有 2 红 3 绿 5 蓝球，有放回地摸 6 次，摸出 2 红 1 绿 3 蓝的概率 $P(X_1 = 2, X_2 = 1, X_3 = 3) = \dfrac{6!}{2!\ 1!\ 3!} \cdot 0.2^2 \cdot 0.1^1 \cdot 0.3^3$.

†：多项式系数：源于多项式展开 $(x_1 + x_2 + \dots + x_k)^n = \sum_{\sum n_i = n} \binom{n}{n_1\ n_2\ \dots\ n_k} x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}$，其计算方法是 $\binom{n}{n_1\ n_2\ \dots\ n_k} = \binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-\dots-n_{k-1}}{n_k} = \dfrac{n}{n_1!\ n_2!\ \dots\ n_k!}$.

*4. 超幾何分布 H(n,M,N)

- **模型**（无放回摸球）N 个产品，其中 M 个次品，从中任取 n 个. X 为这 n 个中的次品数，则 $X \sim H(n, M, N)$.
- **分布** $P(X = k) = \dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad k = 0, 1, \dots, M$
- **例** 兹有 2 红 3 绿 5 蓝球，无放回地摸 6 次，摸出 2 红 1 绿 3 蓝的概率 $P(X_1 = 2, X_2 = 1, X_3 = 3) = \dfrac{\binom{2}{2}\binom{3}{1}\binom{5}{3}}{\binom{10}{6}}$.

二项分布与超几何分布的联系

若 $\lim_{N\to\infty} \frac{M}{N} = p \in (0, 1)$，则

$$\lim_{N\to\infty} \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}} = \binom{n}{m} p^m (1-p)^{n-m}$$

当总量 N 足够大时，如果次品率 $\frac{M}{N}$ 趋向于一个常数 p，无放回地进行抽取（超几何分布）近似为有放回地抽取（二项分布）.

5. 泊松分布

- **模型** X 为某个随机事件发生的次数，假设每次事件发生与否相互独立，且平均事件发生 λ 次，则 $X \sim \pi(\lambda)$.
- **分布** $\boxed{P(X=k)=\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}}, \quad k=0,1,2,\ldots$
- **期望和方差** $\mathrm{E}X=\lambda$，$\mathrm{D}X=\lambda$

独立条件下泊松分布对于其中的参数 λ 具有可加性．即 $X_1 \sim \pi(\lambda_1), X_2 \sim \pi(\lambda_2)$，$X_1, X_2$ 独立，则有 $X_1+X_2 \sim \pi(\lambda_1+\lambda_2)$.

泊松定理（泊松分布与二项分布的联系）

若 $\lambda=\lim\limits_{n\to\infty} np_n>0$，

$$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k e^{-\lambda}}{k!}$$

当 n 足够大，且此时平均发生的事件数趋向于常数 $\lambda=np_n$ 时，二项分布 $b(n,p_n)$ 实质上就是泊松分布 $\pi(np_n)$，此时二项分布的概率值可以用泊松分布的概率值近似.

*6. 几何分布

- **模型** 重复进行随机事件，直到事件发生为止才停下，X 为首次发生时共做的事件的次数，每次发生的概率均为 p，则 $X \sim G(p)$.
- **分布** $P(X=k)=(1-p)^{k-1}p, \quad k=1,2,\ldots, \quad 0<p<1$
- **期望和方差** $\mathrm{E}X=\frac{1}{p}$，$\mathrm{D}X=\frac{q}{p^2}$

具有 **无记忆性**，即 $P(X=m+k \mid X>m)=P(X=k), \quad \forall m,k=1,2,\ldots$，这是几何分布的等价条件.

*7. 帕斯卡分布

- **模型** 重复进行随机事件，直到发生 r 次为止才停止．X 为到停止为止时事件发生与未发生的次数之和，事件每次发生的概率为 p，则 $X \sim \mathrm{Pas}(r, p)$．
- **分布** $P(X=k)=\binom{k-1}{r-1}(1-p)^{k-r}p^r, \quad k=r, r+1, \ldots$
- **期望和方差** $\mathrm{E}X=\dfrac{r}{p}$，$\mathrm{D}X=\dfrac{rq}{p^2}$

当 $r=1$ 时，即为 $G(p)$．
一个服从 $\mathrm{Pas}(r, p)$ 的随机变量为 r 个相互独立的服从 $G(p)$ 的随机变量之和．在独立的条件下帕斯卡分布对于参数 r 具有可加性．

§2.2 連續型隨機變量

随机变量 X 有不可列个不同取值，如一个区间．其在任意单点处取值的概率为 0，故引入概率密度函数．

§2.2.1 連續型隨機變量的分布

连续型随机变量 X 的 **概率分布函数 $F(x)$** 和 **概率密度函数 $f(x)$** 满足

$$\begin{cases} f(x)=F'(x) \\ P(X\leqslant x)=F(x)=\displaystyle\int_{-\infty}^{x} f(x)\mathrm{d}x \\ P(a<X\leqslant b)=F(b)-F(a)=\displaystyle\int_{a}^{b} f(x)\mathrm{d}x \end{cases}$$

1. $F(x)$ 右连续，不严格单调递增；
2. $f(x)\geqslant 0$；
3. $F(+\infty)=\displaystyle\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1,\ F(-\infty)=0$.
4. 利用定积分第一中值定理知 $\displaystyle\lim_{\varepsilon\to 0}\int_{a-\varepsilon}^{a+\varepsilon} f(x)\mathrm{d}x=0$，即连续型随机变量在任意单点处取值的概率为 0，因此，上面的 $P(a<X\leqslant b)$ 中两个等号可任意取．

§2.2.2 連續型隨機變量的函數的分布

已知连续型随机变量 X 的概率分布，$Y=g(X)$，欲求 Y 的概率分布．此时，不能够效仿离散型时的情况，即 $P(Y=k)=f_X(x)$ 并不成立[†]．应先求出分布函数，回到概率的式子，再根据分布函数求出密度．

[†]：问题在于，连续型随机变量在任意单点处取值的概率均为 0，而密度关注的是随机变量在一个区间中取值的概率，而非一点．

例 兹有 $f_X(x)=\begin{cases}1, & 0<x<1\\ 0 & \text{otherwise}\end{cases}$，且 $Y=\mathrm{e}^X$，求 Y 的概率密度函数．

解 当 $y\leqslant 0$ 时，x 无解，此时 $f_Y(y)=0$，以下研究 $y>0$．

$$F_Y(y)=P(Y\leqslant y)=P(\mathrm{e}^X\leqslant y)=P(X\leqslant \ln y)=F_X(\ln y)$$

将上式对 y 求导，有

$$f_Y(y)=f_X(\ln y)\cdot\frac{1}{y}$$

因此，

$$f_Y(y)=\begin{cases}1\cdot\dfrac{1}{y}, & 0<\ln y<1\\ 0\cdot\dfrac{1}{y}, & \text{otherwise}\end{cases}$$

即

$$f_Y(y)=\begin{cases}\dfrac{1}{y}, & 1<y<\mathrm{e}\\ 0, & \text{otherwise}\end{cases}$$

§2.2.3 常見連續概率分布

为方便，仅在概率密度函数中写出 $f(x) \neq 0$ 的情况，即省略了 $f(x)=0, \quad \text{otherwise}$.

1. 均匀分布

- **记号** $X \sim U(a,b), \quad a<b$
- **密度** $f(x)=\dfrac{1}{b-a}, \quad x \in (a,b)$
- **期望和方差** $\mathrm{E}X=\dfrac{a+b}{2}$，$\mathrm{D}X=\dfrac{(b-a)^2}{12}$
- **应用** 产生随机数.

2. 指数分布

- **记号** $X \sim \varepsilon(\lambda), \quad \lambda>0$（或 X 服从参数为 θ $(\theta>0)$ 的指数分布）
- **密度** $f(x)=\lambda \mathrm{e}^{-\lambda x}, \quad x>0$（或 $f(x)=\dfrac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}, \quad x>0$）
- **期望和方差** $\mathrm{E}X=\lambda^{-1}=\theta$，$\mathrm{D}X=\lambda^{-2}=\theta^2$
- **应用** 相邻两次随机事件发生之间的等待时间的分布、电子元件的寿命分布.

具有 **无记忆性**，即 $\forall s,t>0, P(X>s+t \mid X>s)=P(X>t)$，这是指数分布的等价条件.

*3. 伽馬分布

- **记号** $X \sim \Gamma(\alpha, \beta) \quad \alpha, \beta > 0$
- **应用** 气象学中一段时间之内的降水量分布.
- **期望和方差** $\mathrm{E}X = \alpha\beta^{-1}$，$\mathrm{D}X = \alpha\beta{-2}$
- **密度** $f(x) = \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\mathrm{e}^{-\beta x}, \quad x > 0$，其中伽马函数 $\Gamma(\alpha) = \displaystyle\int_0^\infty x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x$.

当 $\alpha = 1$ 时，即为指数分布. 在独立的条件下伽马分布关于参数 α 具有可加性，即服从伽马分布的随机变量可以看成 α 个服从指数分布 $\varepsilon(\beta)$ 的独立随机变量之和.

4. 正態分布

- **记号** $X \sim N(\mu, \sigma^2), \quad \sigma > 0$
- **密度** $f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$
- **期望和方差** $\mathrm{E}X = \mu$，$\mathrm{D}X = \sigma^2$

特别地，**标准正态分布** $N(0,1)$ 的密度 $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$，分布函数

$$\boxed{\Phi(a) = \int_{-\infty}^{a} \varphi(x)\mathrm{d}x}$$

满足 $\forall x$, $\Phi(x) + \Phi(-x) = 1$，且 $X \sim N(\mu, \sigma^2)$ 时，

$$\boxed{P(a < X < b) = \Phi\Big(\frac{b-\mu}{\sigma}\Big) - \Phi\Big(\frac{a-\mu}{\sigma}\Big)}$$

这统一化求正态分布概率的过程，依标准正态分布的概率即可求出所有不同正态分布的概率.

例 兹有 $X \sim N(0,1)$，$Y = 2X^2 + 1$，求 $f_Y(y)$.

解 对 $y > 1$，

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P(2X^2 + 1 \leqslant y) \\ &= P\Big\{-\sqrt{\frac{y-1}{2}} \leqslant X \leqslant \sqrt{\frac{y-1}{2}}\Big\} \\ &= \Phi\Big(\sqrt{\frac{y-1}{2}}\Big) - \Phi\Big(-\sqrt{\frac{y-1}{2}}\Big) \\ &= 2\Phi\Big(\sqrt{\frac{y-1}{2}}\Big) - 1 \end{aligned}$$

将上式对 y 求导，有

$$f_Y(y) = 2\frac{\mathrm{d}}{\mathrm{d}y}\Phi\Big(\sqrt{\frac{y-1}{2}}\Big) = 2\varphi\Big(\sqrt{\frac{y-1}{2}}\Big)\frac{\mathrm{d}}{\mathrm{d}y}\sqrt{\frac{y-1}{2}} = \cdots$$

因此，

$$f_Y(y) = \begin{cases} \dfrac{\mathrm{e}^{-\frac{y-1}{4}}}{2\sqrt{2(y-1)}}, & y > 1 \\ 0, & \text{otherwise} \end{cases}$$

§3 隨機向量與分布

若 X,Y 为随机变量，则随机向量 (X,Y) 的 **联合分布函数**

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

其性质可由一维情况自然地推广.

若是只关心 X 的概率，故遍历 Y 以消除它的影响，有关于 X 的 **边缘分布函数**

$$F_X=\lim_{y\to+\infty}F(x,y)$$

注意 $\lim\limits_{y\to+\infty}P(X\leqslant x,Y\leqslant y)=P(X\leqslant x)$.

§3.1 離散型隨機向量

对于离散型随机向量 (X,Y):

§3.1.1 離散型隨機向量的分布

概率分布 联合概率分布表 $P(X=x_i,Y=y_j)=p_{ij}$，满足 $0\leqslant p_{ij}\leqslant 1$，$\sum_{i,j}p_{ij}=1$.

边缘分布 $P(X=x_i)=\sum_j p_{ij}$，$P(Y=y_j)=\sum_i p_{ij}$，即表中按行或列求和.

§3.1.2 離散型隨機向量的函數的分布

已知 (X,Y) 的联合分布，为求 $Z=g(X,Y)$ 的分布，类似一维，只需合并分布表.

§3.1.3 離散型隨機向量的條件分布

已知 $Y=y_j$ 条件下，X 的条件概率分布

$$P(X=x_i \mid Y=y_j)=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{\sum_i p_{ij}}$$

§3.1.4 離散型隨機向量的分量獨立性

$$\begin{aligned} X \text{ 与 } Y \text{ 独立} &\iff F(x,y)=F_X(x)F_Y(y) \\ &\iff \forall i,j,\ \ P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j) \end{aligned}$$

只要有一者不满足，X 与 Y 就不独立.

§3.2 連續型隨機向量

§3.2.1 連續型隨機向量的分布

连续型随机向量[†] (X,Y) 的 **联合分布函数 $\boldsymbol{F(x,y)}$** 和 **联合密度函数 $\boldsymbol{f(x,y)}$** 满足

$$\begin{cases} f(x,y)=\dfrac{\partial^2}{\partial x\partial y}F(x,y)^{\ddagger} \\ P(X\leqslant x,Y\leqslant y)=F(x,y)=\displaystyle\int_{-\infty}^{y}\int_{-\infty}^{x}f(x,y)\mathrm{d}x\mathrm{d}y \end{cases}$$

† : 连续型随机向量的定义是由存在联合密度函数给出的．需要注意的是，即使 X,Y 均为连续型随机变量，(X,Y) 也不一定是连续型随机向量，需要验证 $\displaystyle\iint_{\mathbb{R}^2}f(x,y)\mathrm{d}x\mathrm{d}y=1$．例如连续型随机变量 $X,Y\sim U(0,1)$，容易得到 (X,Y) 的联合密度恒为 0，故 (X,Y) 不是连续型随机向量.

‡ : 若 $\dfrac{\partial^2}{\partial x\partial y}F(x,y)$ 不存在，则 $f(x,y)$ 取 0.

关于 X 的 **边缘密度函数 $\boldsymbol{f_X(x,y)}$**

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

§3.2.2 連續型隨機向量的函數的分布

已知连续型随机向量 (X,Y) 的概率分布，$Z=g(X,Y)$，欲求 Z 的概率分布，类似地，先求分布函数.

大概知道推导方法即可，这些结果无需记忆.

若 $Z=X+Y$，

$$
\begin{aligned}
F_{X+Y}(z)=P(X+Y\leqslant z) &= \iint_{x+y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{z-x} f(x,y)\mathrm{d}y \\
&= \int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{z} f(x,z-x)\mathrm{d}z \\
&= \int_{-\infty}^{z}\mathrm{d}z\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x
\end{aligned}
$$

故

$$
f_{X+Y}(z)=F'_{X+Y}(z)=\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x
$$

特别地，若 X,Y 独立，

$$
f_{X+Y}(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x
$$

若 $Z=X-Y$，则 $f_{X-Y}(z)=\displaystyle\int_{-\infty}^{+\infty} f(x,x-z)\mathrm{d}x$

若 $Z=\dfrac{Y}{X}$，则 $f_{Y/X}(z)=\displaystyle\int_{-\infty}^{+\infty} |x|f(x,xz)\mathrm{d}x$

若 $Z=XY$，则 $f_{XY}(z)=\displaystyle\int_{-\infty}^{+\infty} \left|\frac{1}{x}\right|f\left(x,\frac{z}{x}\right)\mathrm{d}x$

若 $Z=\max(X,Y)$，则 $F_{\max}(z)=F_X(z)F_Y(z)$

若 $Z=\min(X,Y)$，则 $F_{\min}(z)=1-[1-F_X(z)][1-F_Y(z)]$

例 兹有 $f(x,y)=\begin{cases}x+y, & 0<x<1,0<y<1\\ 0 & \text{otherwise}\end{cases}$，且 $Z=X+Y$，求 Z 的概率密度函数.

析 大致的形式是 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x$，需对字母的范围进行讨论.

解 由 $z=x+y$, $0<x<1$, $0<y<1$，得 $0<z<2$.

画出 $x=x(z)$ 的图形，确定 z 的讨论边界.

当 $0<z<1$ 时，$0<x<z$, $0<z-x<1$，此时

$$f_Z(z)=\int_0^z f(x,z-x)\mathrm{d}x=\int_0^z (x+z-x)\mathrm{d}x=z^2$$

当 $1\leqslant z<2$ 时，$z-1<x<1$, $0<z-x<1$，此时

$$f_Z(z)=\int_{z-1}^1 f(x,z-x)\mathrm{d}x=2z-z^2$$

故

$$f_Z(z)=\begin{cases}z^2, & 0<z<1\\ 2z-z^2, & 1\leqslant z<2\\ 0, & \text{otherwise}\end{cases}$$

§3.2.3 連續型隨機向量的條件分布

已知 $Y=y$ 条件下，X 的条件概率分布

$$\boxed{F_{X|Y}(x\mid y)=P(X\leqslant x,Y=y)=\int_{-\infty}^x \frac{f(x,y)}{f_Y(y)}\mathrm{d}x}$$

条件概率密度

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

§3.2.4 連續型隨機向量的分量獨立性

$$\begin{aligned} X \text{ 与 } Y \text{ 独立} &\iff F(x,y)=F_X(x)F_Y(y) \\ &\iff f(x,y)=f_X(x)f_Y(y) \end{aligned}$$

例 10 个变量相互独立，且服从 $N(0,1)$，求联合概率密度.

解 $f(x_1,x_2,\ldots,x_{10})=\prod\limits_{i=1}^{10}\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$, $\quad x_i\in\mathbb{R}$，这个乘积的结果是由独立性保证的.

§3.2.5 常見二維連續概率分布

*1. 均匀分布

(X,Y) 在 D 上均匀分布，则密度 $f(x,y)=\dfrac{1}{m(D)}$, $\quad (X,Y)\in D$，其中 $D\subset\mathbb{R}^2$，面积 $m(D)$ 有限.

例 (X,Y) 在单位圆内部 D 上均匀分布，则密度
$f(x,y)=\dfrac{1}{\pi}$, $\quad (X,Y)\in D$. 关于 X 的边缘密度为

$$f_X(x)=\int_y f(x,y)\mathrm{d}y=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}\mathrm{d}y=\frac{2}{\pi}\sqrt{1-x^2},\quad -1\leqslant x\leqslant 1$$

同理，$f_Y(y)=\dfrac{2}{\pi}\sqrt{1-y^2}$, $\quad -1\leqslant y\leqslant 1$. 因 $f_X(x)f_Y(y)\neq f(x,y)$，故 X,Y 不独立.

2. 二維正態分布

若 $(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$，则 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$.

在给定 $X=x$ 条件下 Y 的条件分布仍旧为正态分布.

§4　随機變量的數字特征

§4.1　數學期望

§4.1.1　數學期望的定義和計算

离散型随机变量

离散型随机变量 X 概率分布为 $P(X=x_j)=p_j$，若 $\sum\limits_{j=1}^{+\infty} p_j\left|x_j\right|$ 收敛 †，则

$$\boxed{\mu=\mathrm{E}X=\sum_{j=1}^{+\infty} p_j x_j}$$

即是高中时所学的，将分布表上下两行对应相乘再求和.

连续型随机变量

连续型随机变量 X 概率密度为 $f(x)$，若 $\displaystyle\int_{-\infty}^{+\infty} f(x)|x|\mathrm{d}x$ 收敛，则

$$\boxed{\mu=\mathrm{E}X=\int_{-\infty}^{+\infty} x f(x)\mathrm{d}x}$$

† : 即 $\sum\limits_{j=1}^{+\infty} p_j x_j$ 绝对收敛. 希望这个级数的求和中任意两项的交换不影响级数的敛散性，也不影响级数之值且绝对收敛的级数具有交换律 ‡，故要求级数/积分绝对收敛.

其函数的计算方法，如 $\mathrm{E}X^2=\displaystyle\int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x$，

$\mathrm{E}g(X)=\displaystyle\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$，.

§4.1.2　數學期望的性質

1. $\mathrm{E}(a+bX+cY)=a+b\mathrm{E}X+c\mathrm{E}Y$（线性性质）

2. $X_1, \ldots, X_n$ **相互独立** 且数学期望均存在，则 $\mathrm{E} \prod_{i=1}^{n} X_i = \prod_{i=1}^{n} \mathrm{E} X_i$

§4.2 方差

§4.2.1 方差的定義和計算

若随机变量 X 数学期望 μ 存在且有限，则 **方差** σ_{XX} 和 **标准差** σ_X

$$\boxed{\sigma_{XX} = \sigma_X^2 = \mathrm{D}X = \mathrm{E}(X - \mathrm{E}X)^2 = \mathrm{E}X^2 - \mathrm{E}^2 X}$$

特别地，**标准化变量** $X^* = \dfrac{X-\mu}{\sigma}$ 的期望 $\mathrm{E}X^* = 0$，方差 $\mathrm{D}X^* = 1$.

§4.2.2 方差的性質

1. $\boxed{\mathrm{D}(a + bX) = b^2 \mathrm{D}X}$
2. $\mathrm{D}X = \mathrm{E}(X-\varepsilon)^2 - (\mu - \varepsilon)^2$，若 $\varepsilon \neq \mu$，有 $\mathrm{D}X < E(X-\varepsilon)^2$.
   这表明，方差当且仅当随机变量取到数学期望时取到最小值.
3. $\boxed{\forall \varepsilon > 0, P\big\{|X - \mu| \geqslant \varepsilon\big\} \leqslant \dfrac{\sigma^2}{\varepsilon^2}}$（Chebyshev 亲亲不等式）
4. $\mathrm{D}X = 0 \iff P\big\{X = \mu\big\} = 1$.
5. $\mathrm{D}\sum\limits_{j=1}^{n} X_j = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(\mathrm{E}X_iX_j - \mathrm{E}X_i\mathrm{E}X_j)$.
   特别地，$\mathrm{D}(X+Y) = \mathrm{D}X + \mathrm{D}Y + 2\,\mathrm{Cov}(X,Y)$;
   特别地，若 $X_1, \ldots, X_n$ 相互独立，$\mathrm{D}\sum\limits_{j=1}^{n} X_j = \sum\limits_{j=1}^{n} \mathrm{D}X_j$.

例 设 $X \sim U(0,6)$, $Y \sim b(12,4)$，且 X, Y 相互独立，用 Chebyshev 不等式估计 $P\big\{X - 3 < Y < X + 3\big\}$.

解 所求即 $P\big\{|X-Y| < 3\big\}$. 因 $\mathrm{E}(X-Y) = \mathrm{E}X - \mathrm{E}Y = 3 - 3 = 0$，且由独立性知 $\mathrm{D}(X-Y) = \mathrm{D}X + \mathrm{D}Y = 3 + \dfrac{9}{4} = \dfrac{21}{4}$,

故原式 $= P\big\{|(X-Y) - \mathrm{E}(X-Y)| < 3\big\} \geqslant 1 - \dfrac{\mathrm{D}(X-Y)}{3^2} = \dfrac{5}{12}$.

§4.3　協方差、相關系數

§4.3.1　協方差、相關系數

随机变量 X,Y 的 **协方差 $\boldsymbol{\sigma_{XY}}$**（或 $\mathrm{Cov}(X,Y)$）

$$\boxed{\sigma_{XY}=\mathrm{E}(X-\mathrm{E}X)(Y-\mathrm{E}Y)=\mathrm{E}XY-\mathrm{E}X\mathrm{E}Y}$$

相关系数 $\boldsymbol{\rho_{XY}}$

$$\boxed{\rho_{XY}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}}$$

$|\rho_{XY}|\leqslant 1$，等号成立当且仅当 $\exists a,b,\ \text{s.t.}\ P(a+bX)=1$.

§4.3.2　獨立性、相關性

协方差描述了两个变量之间的线性关系，相关系数的正负决定了两个随机变量是正相关还是负相关的，也即描述了在变化过程中，两个变量是同方向变化，还是反方向变化，以及同向或反向程度如何．相关系数是剔除了两个变量量纲影响，标准化后的特殊协方差.

因此，协方差和相关系数可用于判断独立性和相关性．注意，独立则不相关，不相关却不一定独立.

- X,Y 独立 $\Longrightarrow \mathrm{E}XY=\mathrm{E}X\mathrm{E}Y \Longrightarrow \sigma_{XY}=\rho_{XY}=0$
- X,Y 不相关 $\iff \sigma_{XY}=\rho_{XY}=0$

數理統計部分

§5 統計推斷緒

§5.0.1 大數定律

——当试验次数足够大的时候，样本平均值的统计量会和期望无限接近期望，即频率无限接近于概率.

大数定律 取来自于总体 X 的相互独立的样本 $X_1, X_2, \cdots, X_n$，对于 $\forall \varepsilon > 0$，有 $\lim\limits_{n\to\infty} P\{|\overline{X} - \mu| < \varepsilon\} = 1$，其中 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 是样本平均值，此即 **弱大数定律 (Khinchin 大数定律)**. 特别地，设 n_A 为 n 次独立重复试验中事件 A 发生的次数，p 为事件 A 在每次试验中发生的概率，对于 $\forall \varepsilon > 0$，有 $\lim\limits_{n\to\infty} P\{|\frac{n_A}{n} - p| < \varepsilon\} = 1$，此即 **Bernoulli 大数定律**.

§5.0.2 中心極限定理

——大量的相互独立的随机因素的综合影响形成的结果往往近似的服从正态分布.

中心极限定理 取来自于总体 X 的相互独立的样本 $X_1, X_2, \cdots, X_n$，当 $n \to \infty$ 时，它们的和 $\sum\limits_{k=1}^{n} X_k$ 的标准化变量

$$\boxed{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \sim N(0,1)}$$

即分布函数 $F_n(x)$ 满足 $\lim\limits_{n\to\infty} F_n(x) = \Phi(x)$. 特别地，二项分布的中心极限定理称为 **De Moivre-Laplace 定理**.

例 兹有某种物品的寿命 $X \sim \varepsilon(100)$，随机地取相互独立的 16 件，求它们的寿命之和大于 1920 的概率.

解 设 $X_i, i = 1, 2, \ldots, 16$，记 $Y = \sum\limits_{i=1}^{n} X_i$，则 $\frac{Y - 6\times 100}{\sqrt{16}\times 100}$ 近似地服从 $N(0,1)$，因此 $P(Y > 1920) = 1 - \Phi(\frac{1920-1600}{400}) = 0.2119$.

§5.1 樣本與統計量

——从总体抽取样本，观测样本得到样本值，样本及它们的函数是统计推断的依据.

统计量 来自于总体 X 的样本 $X_1, X_2, \cdots, X_n$ 的不含未知参数的函数 $g(X_1, X_2, \cdots, X_n)$，如

$$\text{样本平均值} \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\text{样本方差} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)$$

$$\text{样本标准差} \quad S = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)}$$

$$k\ \text{阶原点矩} \quad A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \quad k = 1, 2, \cdots$$

$$k\ \text{阶中心矩} \quad B_k = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^k, \quad k = 2, 3, \cdots$$

这些样本的观察值（就是测出来的数值）是 $x_1, x_2, \cdots, x_n$，上述统计量的观察值便是 $g(x_1, x_2, \cdots, x_n)$，如

$$\text{样本平均值的观察值} \quad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\text{样本方差的观察值} \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right)$$

$$\cdots \qquad \cdots$$

常用性质：$\mathrm{E}\overline{X} = \mathrm{E}X$, $\mathrm{D}\overline{X} = \dfrac{1}{n}\mathrm{D}X$, $\mathrm{E}S^2 = \mathrm{D}X$.

§5.2 三大抽样分布

——使用抽样统计量描述样本的分布.

1. 卡方分布

- **定义** 来自于总体 X 的样本 $X_1, X_2, \cdots, X_n$ 的统计量 $\chi^2 = \sum\limits_{i=1}^{n} X_i^2$ 服从自由度为 n 的 χ^2 分布，即 $\chi^2(n)$
- **图形** 分布于第一象限内，呈右偏态，随着自由度 n 的增大，峰值向右下方移动，趋近于正态分布
- **上 $\boldsymbol{\alpha}$ 分位点性质** 无
- **期望和方差** $\mathrm{E}\chi^2 = n$，$\mathrm{D}\chi^2 = 2n$
- **可加性** 设 $\chi_1^2 \sim \chi^2(n_1), \chi_2^2 \sim \chi^2(n_2)$ 且 χ_1^2, χ_2^2 相互独立，则 $\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$

2. t 分布

- **定义** 设 $X \sim N(0,1), Y \sim \chi^2(n)$，且 X, Y 相互独立，则 $t = \frac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布，即 $t(n)$
- **图形** 与正态分布类似的单峰偶函数，自由度 n 越大，峰值越大越聚集，曲线越接近标准正态分布曲线
- **上 $\boldsymbol{\alpha}$ 分位点性质** $t_{1-\alpha}(n) = -t_\alpha(n)$

3. F 分布

- **定义** 设 $U \sim \chi^2(n_1), V \sim \chi^2(n_2)$，且 U, V 相互独立，则 $F = \frac{U/n_1}{V/n_2}$ 服从自由度为 (n_1, n_2) 的 F 分布，即 $F \sim F(n_1, n_2)$
- **性质** $F \sim F(n_1, n_2) \Longrightarrow \dfrac{1}{F} \sim F(n_2, n_1)$
- **图形** 与卡方分布类似的右偏单峰曲线，分布于第一象限内
- **上 $\boldsymbol{\alpha}$ 分位点性质** $F_{1-\alpha}(n_1, n_2) = \dfrac{1}{F_\alpha(n_2, n_1)}$

§5.3　四大定理

——依据正态总体的样本均值与样本方差的分布分析随机取样得到的数据，大规模的统计数据都近似地符合正态分布，造就了这四大定理的绝对的奠基地位.

引理 $\begin{cases} \mathrm{E}\overline{X} = \mu \\ \mathrm{D}\overline{X} = \dfrac{\sigma^2}{n} \end{cases}$

设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $N(\mu, \sigma^2)$ 的样本，样本均值和方差分别为 $\overline{X}, S^2$，

定理一 $\overline{X} \sim N(\mu, \sigma^2/n)$　i.e. $\boxed{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)}$

定理二 $\boxed{\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)}$

定理三 $\boxed{\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)}$

设 $X_1, X_2, \cdots, X_{n_1}$ 与 $Y_1, Y_2, \cdots, Y_{n_2}$ 分别为来自于正态分布总体 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ 相互独立的样本，均值 $\overline{X}, \overline{Y}$，方差 S_1^2, S_2^2，

定理四之一 $\dfrac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$

定理四之二 $\sigma_1^2 = \sigma_2^2 = \sigma^2 \Longrightarrow \dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$

例 兹有随机变量，期望为 5，方差为 0.3，取两组容量为 80 的样本，这两组的均值分别为 X, Y，求 $P(-0.1 < \overline{X}-\overline{Y} < 0.1)$.
解 $\mathrm{E}(\overline{X}-\overline{Y}) = \mathrm{E}\overline{X} - \mathrm{E}\overline{Y} = 0$，$\mathrm{D}(\overline{X}-\overline{Y}) = \mathrm{D}\overline{X} + \mathrm{D}\overline{Y} = \frac{0.3}{80} + \frac{0.3}{80} = \frac{3}{400}$. 所求 $= 2\Phi\left(\frac{0.1-0}{\sqrt{3/400}}\right) - 1 = 0.7098$.

例 兹有正态分布总体，从中抽容量为 16 的样本，求 $P\left\{\dfrac{S^2}{\sigma^2} \leqslant 2.041\right\}$.
解 由定理二，$\dfrac{15S^2}{\sigma^2} \sim \chi^2(15)$，所求 $= P\left\{\dfrac{15S^2}{\sigma^2} \leqslant 30.615\right\} = 1 - \left\{\dfrac{15S^2}{\sigma^2} > 30.615\right\}$，查表得 $\chi^2_{0.01}(15) = 30.578$，故所求 $= 1 - 0.01 = 0.99$.

§6 統計推斷之參數估計

§6.1 點估計

——根据观察的样本推测总体，由已知推未知．在已知部分概率分布的情况下，构造一个适当的统计量，把它的观察值作为未知参数的估计值．

§6.1.1 估計量的判斷標準

- **无偏性** $\mathrm{E}\hat{\theta} = \theta$.
- **有效性** $\mathrm{D}\hat{\theta}$ 越小越好.
- **相合性** $\forall\theta \in \Theta, \lim\limits_{n\to\infty} P\left\{|\hat{\theta} - \theta| < \varepsilon\right\} = 1$.

例 确定 c，使 $c\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ 是 σ^2 的无偏估计量.
解 置 $\mathrm{E}c\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \sigma^2$，化简并解得 $c = \dfrac{1}{n-1}$.
注 事实上，此即样本方差 S^2，这也就解释了分母取 $n-1$ 的原因：考虑到无偏性，减小系统偏差，希望 $\mathrm{E}S^2 = \sigma^2$.

§6.1.2 矩估計法

——矩估计已死.

样本 k 阶矩 $A_k = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^k$，依中心极限定理，其概率收敛为 μ_k. 若有 k 个未知参数，将 $1, 2, \ldots, k$ 阶矩 $\mu_1, \mu_2, \ldots, \mu_k$ 作为估计量并用未知参数表示出来，然后用已知的样本统计量 $A_1, A_2, \ldots, A_k$ 去替换，解出这些未知参数.

§6.1.3 極大似然估計法

——直观地，总体在已知的样本上发生的概率较大.

为了求 MLE (Memory Limit Exceeded，指大学牲的脑子已经装不下新知识了)，只需在期末周前高强度学习，直到不能再学习为止，统计已经学会的知识，即是 **大学牲死然估计**.

事实上，数理统计中 MLE 指极大似然估计 (Maximum Likelihood Estimate)，首先定义极大似然函数

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

将 $L(\theta)$ 取到最大值的 $\hat{\theta}$ 值作为 θ 的极大似然估计量. 为了求出 $\hat{\theta}$，一般置 $\dfrac{\mathrm{d}}{\mathrm{d}\theta} \ln L(\theta) = 0$.

例 设 $X_1, X_2, \cdots, X_n$ 为来自总体的样本，$x_1, x_2, \cdots, x_n$ 为相应的样本值，求总体概率密度或分布中未知参数的矩估计量和矩估计值，以及极大似然估计量和极大似然估计值. 兹有总体 $f(x) = \begin{cases} \theta c^\theta x^{-(\theta+1)}, & x > c \\ 0, & \text{otherwise} \end{cases}$，其中 $c > 0$ 为已知，$\theta > 1$ 为未知参数.

解 矩估计:

$\mu_1 = \mathrm{E}X = \displaystyle\int_{-\infty}^{+\infty} x f(x) \mathrm{d}x = \int_{c}^{+\infty} \theta c^\theta x^{-\theta} \mathrm{d}x = \frac{c\theta}{\theta - 1}$，解得 $\theta = \dfrac{\mu_1}{\mu_1 - c}$，进而 θ 的矩估计量 $\hat{\theta} = \dfrac{A_1}{A_1 - c} = \dfrac{\overline{X}}{\overline{X} - c}$，矩估计值 $\hat{\theta} = \dfrac{\overline{x}}{\overline{x} - c}$.

极大似然估计:

极大似然函数 $L(\theta) = \prod\limits_{i=1}^{n} f(x_i) = \prod\limits_{i=1}^{n} \theta c^\theta x_i^{-(\theta+1)} = \theta^n c^{n\theta} \prod\limits_{i=1}^{n} x_i^{-(\theta+1)}$，置 $\dfrac{\mathrm{d}}{\mathrm{d}\theta} \ln L(\theta) = 0$，得 $\dfrac{\theta}{n} + n \ln c - \sum\limits_{i=1}^{n} \ln x_i = 0$，解得 θ 的极大似然估计值 $\hat{\theta} = \dfrac{n}{\sum_{i=1}^{n} \ln x_i - n \ln c}$，极大似然估计量 $\hat{\theta} = \dfrac{n}{\sum_{i=1}^{n} \ln X_i - n \ln c}$.

评注 如果有多个参量需要估计（P153 例 3、P156 例 5）:

矩估计: $\mu_2 = \mathrm{E}X^2, \mu_k = \mathrm{E}X^k$，解方程组得到未知参数的值;

极大似然估计: 即是多元函数极值，对多个变量分别置 $\dfrac{\mathrm{d}}{\mathrm{d}\theta_i} \ln L(\theta_1, \theta_2, \ldots) = 0$，解方程组.

§6.2 區間估計

§6.2.1 置信區間

——点估计范围太小，又没有一定的系统误差容错范围，结果很不稳定，也不易说服人，故希望找到一个有一定的容错误差的、令人信服的范围，这个范围即是置信区间.

若 $P\{\theta_1 < \theta < \theta_2\} \geqslant 1-\alpha$，称区间 (θ_1, θ_2) 为置信水平为 $1-\alpha$ 的 **置信区间**. 如置信水平为 0.95，表明 100 个样本中，平均只有 5 个不在估计的范围内.

（一）设 $X_1, X_2, \cdots, X_n$ 为来自总体 $N(\mu, \sigma^2)$ 的样本，样本均值与方差 $\overline{X}, S^2$. 在以下条件下，求指定变量的置信水平为 $1-\alpha$ 的区间:

建议根据三大分布四大定理推导，不要硬背结论.

(1) **求 $\boldsymbol{\mu}$ 的置信区间，$\boldsymbol{\sigma^2}$ 已知:**

- 选择定理 **定理一** $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- 置信区间 $P\Big\{-z_{\frac{\alpha}{2}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < z_{\frac{\alpha}{2}}\Big\} = 1-\alpha$
- 化简结果 $\Big(\overline{X} \pm \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\Big)$

(2) **求 $\boldsymbol{\mu}$ 的置信区间，$\boldsymbol{\sigma^2}$ 未知:**

- 选择定理 **定理三** $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$
- 置信区间 $P\Big\{-t_{\frac{\alpha}{2}}(n-1) < \frac{\overline{X}-\mu}{S/\sqrt{n}} < t_{\frac{\alpha}{2}}(n-1)\Big\} = 1-\alpha$
- 化简结果 $\Big(\overline{X} \pm \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\Big)$

(3) **求 $\boldsymbol{\sigma^2}$ 的置信区间:**

- 选择定理 **定理二** $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
- 置信区间 $P\Big\{\chi^2_{1-\frac{\alpha}{2}}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}(n-1)\Big\} = 1-\alpha$
- 化简结果 $\Big(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\Big)$

（二）两个总体 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$，求指定变量的置信水平为 $1-\alpha$ 的区间:

(1) **求两个总体均值差 $\boldsymbol{\mu_1-\mu_2}$ 的置信区间，方差 $\boldsymbol{\sigma_1^2,\sigma_2^2}$ 均已知**:

- 选择定理 **定理一** $\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$
- 化简结果 $\left((\overline{X}-\overline{Y})\pm z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$

(2) **求两个总体均值差 $\boldsymbol{\mu_1-\mu_2}$ 的置信区间，方差 $\boldsymbol{\sigma_1^2=\sigma_2^2=\sigma^2}$ 但未知**:

- 选择定理 **定理四之二** $\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$
- 置信区间 $P\left\{-t_{\frac{\alpha}{2}}(n_1+n_2-2)<\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}<t_{\frac{\alpha}{2}}(n_1+n_2-2)\right\}=1-\alpha$
- 化简结果 $\left((\overline{X}-\overline{Y})\pm t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$

(3) **求两个总体方差比 $\boldsymbol{\sigma_1^2/\sigma_2^2}$ 的置信区间**:

- 选择定理 **定理四之一** $\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$
- 置信区间 $P\left\{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)<\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}<F_{\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}=1-\alpha$
- 化简结果 $\left(\frac{S_1^2}{S_2^2}\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right)$

§6.2.2　單側置信區間

> ——不关心上下限，而是某一个临界
>
> 单侧置信下限 θ_1 定义为 $P\{\theta>\theta_1\}\geqslant 1-\alpha$，单侧置信上限 θ_2 定义为 $P\{\theta<\theta_2\}\geqslant 1-\alpha$.

将前一个部分中的参数 $\frac{\alpha}{2}$ 改为 α，相应的区间上/下界即是单侧置信上/下限.

§7 統計推斷之假設檢驗

§7.1 假設檢驗

§7.1.1 雙邊假設檢驗

——对于一个未知的总体，根据样本提出 **零假设** H_0 及它反方向的假设即 **备择假设** H_1，并做出决策，是接受还是拒绝这个假设．

第 I 类错误 在假设 H_0 为真的时候拒绝了它，在做检验的时候需要控制这一类错误．
第 II 类错误 H_1 为真的时候接受了 H_0．一般不考虑，因为在控制第 I 类错误发生概率时，就已经间接的提高了第 II 类错误的发生概率，而降低它的方法只有提高样本容量．
显著性检验 只考虑第 I 类错误，不考虑第 II 类错误．拒绝 H_0 表明估计值和总体的这个参数值的差异显著．
显著性水平 第 I 类错误的最大概率 α．
拒绝域 取到使 H_0 被拒绝的值的值域，如果观测值落在拒绝域内，则在 α 这个显著性水平下拒绝 H_0，否则接受 H_0．

$$\boxed{\text{置信水平为 } 1-\alpha \text{ 的置信区间} \iff \text{显著性水平为 } \alpha \text{ 的假设检验的接受域}}$$

§7.1.2 單邊假設檢驗

——单方向的假设．设 x 是需要检验的变量，定义

- **右边检验** $H_0 : x \leqslant x_0$, $H_1 : x > x_0$;
- **左边检验** $H_0 : x \geqslant x_0$, $H_1 : x < x_0$.

$$\boxed{\begin{array}{l}\text{单侧置信下限对应的区间} \iff \text{左边假设检验的接受域}\\ \text{单侧置信上限对应的区间} \iff \text{右边假设检验的接受域}\end{array}}$$

下面举一些例子：

取显著性水平为 α：

建议依据置信区间推导，不要硬背结论.

(1) **单个总体均值，且 $\boldsymbol{\sigma^2}$ 已知**：假设 $H_0:\mu=\mu_0$，$H_1:\mu\neq\mu_0$，

- 检验法：**Z 检验法**，$Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$
- 拒绝域：$|z|\geqslant z_{\frac{\alpha}{2}}$
- 右边检验：假设 $H_0:\mu\leqslant\mu_0$，$H_1:\mu>\mu_0$，拒绝域 $z\geqslant z_\alpha$；
- 左边检验：假设 $H_0:\mu\geqslant\mu_0$，$H_1:\mu<\mu_0$，拒绝域 $z\leqslant -z_\alpha$；

(2) **单个总体均值，且 $\boldsymbol{\sigma^2}$ 未知**：假设 $H_0:\mu=\mu_0$，$H_1:\mu\neq\mu_0$，

- 检验法：**t 检验法**，$t=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$
- 拒绝域：$|t|\geqslant t_{\frac{\alpha}{2}}(n-1)$

(3) **单个总体方差**：假设 $H_0:\sigma^2=\sigma_0^2$，$H_1:\sigma^2\neq\sigma_0^2$，

- 检验法：**卡方检验法**，$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}$
- 拒绝域：$\chi^2\leqslant\chi^2_{1-\frac{\alpha}{2}}(n-1)\ \vee\ \chi^2\geqslant\chi^2_{\frac{\alpha}{2}}(n-1)$
- 右边检验：假设 $H_0:\sigma^2\leqslant\sigma_0^2$，拒绝域 $\chi^2\geqslant\chi^2_\alpha(n-1)$；
- 左边检验：假设 $H_0:\sigma^2\geqslant\sigma_0^2$，拒绝域 $\chi^2\leqslant\chi^2_{1-\alpha}(n-1)$；

(4) **两个正态总体均值差 $\boldsymbol{\mu_1-\mu_2}$**：在认为 $\sigma_1^2=\sigma_2^2=\sigma^2$ 的前提下 [这需要先检验 (3)]，假设 $H_0:\mu_1-\mu_2=\delta$，

- 检验法：t 检验法，$t=\frac{(\overline{X}-\overline{Y})-\delta}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$
- 拒绝域：$|t|\geqslant t_{\frac{\alpha}{2}}(n_1+n_2-2)$

Jun.13 完结撒花！三天概率论两天数理统计冲刺毕！