概统笔记附录

附录

A.1 常用积分

特殊函数的性质 伽马函数与贝塔函数.

伽马函数与递推式: $\Gamma(x)=\int_0^{+\infty}\mathrm{e}^{-t}t^{x-1}\,\mathrm{d}t=(x-1)\Gamma(x-1)\quad(x>0)$

贝塔函数与关系式: $\mathrm{B}(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t=\dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}\quad(x,y>0)$

勒让德倍量公式: $\Gamma(s)\Gamma\left(s+\dfrac{1}{2}\right)=\dfrac{\sqrt{\pi}}{2^{2n-1}}\Gamma(2s)\quad(s>0)$

余元公式: $B(s,1-s)=\Gamma(s)\Gamma(1-s)=\dfrac{\pi}{\sin\pi s}\quad(0<s<1)$

$$\begin{cases}\Gamma(n)=(n-1)!, & n\in\mathbb{N}^+,\\ \Gamma\left(\dfrac{n}{2}\right)=\dfrac{(n-2)!!}{2^{(n-1)/2}}\sqrt{\pi}, & n\text{ 为正奇数}.\end{cases}$$

$$B(s,s)=\frac{1}{2^{2n-1}}B\left(\frac{1}{2},s\right)\quad(s>0)$$

更多内容可以参考 Euler 积分笔记.

特殊函数的应用

一般的

$$\int_0^1 x^a(1-x^b)^c\,\mathrm{d}x=\frac{1}{b}\mathrm{B}\left(\frac{a+1}{b},c+1\right)\qquad(a>-1,\ b>0,\ c>-1)$$

$$\int_0^{+\infty}\frac{x^a\,\mathrm{d}x}{(1+x^b)^c}=\frac{1}{|b|}\mathrm{B}\left(c-\frac{a+1}{b},\frac{a+1}{b}\right)\qquad\left(\begin{array}{ll}a>-1,\ b>0,\ c>\dfrac{a+1}{b} & \text{或}\\ a<-1,\ b<0,\ c>\dfrac{a+1}{b} & \end{array}\right)$$

$$\int_0^{+\infty}x^n\mathrm{e}^{-ax^p}\,\mathrm{d}x=\frac{\Gamma\left(\dfrac{n+1}{p}\right)}{|p|a^{\frac{n+1}{p}}}\qquad\left(\begin{array}{ll}a>0,\ p>0,\ n>-1 & \text{或}\\ a>0,\ p<0,\ n<-1 & \end{array}\right)$$

特殊的

$$\int_{-\infty}^{+\infty}\frac{x^m\,\mathrm{d}x}{(1+x^2)^n}=\mathrm{B}\left(n-\frac{m+1}{2},\frac{m+1}{2}\right)\qquad(\text{注意积分限})$$
$$\int_0^{+\infty}\frac{x^m\,\mathrm{d}x}{(1+x)^n}=\mathrm{B}\left(n-m-1,m+1\right)$$

$$\int_0^{+\infty}\mathrm{e}^{-ax^p}\,\mathrm{d}x=\frac{\Gamma\left(\dfrac{1}{p}\right)}{pa^{\frac{1}{p}}}$$
$$\int_0^{+\infty}\mathrm{e}^{-x^p}\,\mathrm{d}x=\frac{1}{p}\Gamma\left(\frac{1}{p}\right)$$
$$\int_{-\infty}^{+\infty}\mathrm{e}^{-ax^2}\,\mathrm{d}x=\sqrt{\frac{\pi}{a}}\qquad(\text{注意积分限})$$

$$\int_0^{+\infty} x^n \mathrm{e}^{-ax} = \frac{\Gamma(n+1)}{a^{n+1}} = \frac{n!}{a^{n+1}}$$

$$\int_0^{+\infty} x^n \mathrm{e}^{-ax^2} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2a^{\frac{n+1}{2}}}$$

$$\int_0^{+\infty} x^{2n} \mathrm{e}^{-ax^2}\,\mathrm{d}x = \frac{(2n-1)!!}{2(2a)^n}\sqrt{\frac{\pi}{a}}$$

$$\int_0^{+\infty} x^{2n+1} \mathrm{e}^{-ax^2}\,\mathrm{d}x = \frac{(2n)!!}{(2a)^{n+1}}$$

$$\int_0^{+\infty} x^2 \mathrm{e}^{-ax^2}\,\mathrm{d}x = \frac{1}{4a}\sqrt{\frac{\pi}{a}}$$

有关 Catalan 常数的积分.

A.2 常用分布

定义说明

- 期望 $\mu := E(X)$.
- 方差 $\sigma^2 := E[(X-\mu)^2]$.
- k 阶原点矩 $\alpha_k := E(X^k)$.
- k 阶中心距 $\mu_k := E[(X-\mu)^k]$.
- 偏度系数 $\beta_1 = \dfrac{\mu_3}{\sigma^3}$.
- 峰度系数 $\beta_2 := \dfrac{\mu_4}{\sigma^4} = \dfrac{\mu_4}{\mu_2^2}$.
- 变异系数 $v_c := \dfrac{\sigma}{\mu}$.

A.2.1 一维离散型

1 二项分布

1.1 基础概念

- $X \sim B(n,p)$.
- 理解: 事件发生的概率为 p, 则重复 n 次试验, 事件发生的次数为 x.
- 概率分布: $P(X=i) = b(i;n,p) = \dbinom{n}{i} p^i (1-p)^{n-i}$.

1.2 数字特征

- 最可能数: $x = \lfloor (n+1)p \rfloor$.
- 期望: $E(X) = np$.
- 方差: $\mathrm{Var}(X) = np(1-p)$.
- 母函数: $G(s) = (ps+q)^n$, $s \in (-\infty, +\infty)$.
- 特征函数: $g(t) = (p\mathrm{e}^{\mathrm{i}t} + q)^n$.

1.3 其它性质

- 二项分布和的函数

  $$X_1 \sim B(n_1,p),\ X_2 \sim B(n_2,p) \quad \Rightarrow \quad X_1 + X_2 \sim B(n_1+n_2,p).$$

- 发生偶数次的概率为 $p_n = \dfrac{1}{2}[1+(1-2p)^n]$.
- 记 $f(p) = P(X \le k)$, 则 $f'(p) < 0$, 并且

$$f(p)=\frac{n!}{k!(n-k-1)!}\int_0^{1-p} t^k(1-t)^{n-k-1}\,\mathrm{d}t.$$

1.4 参数估计

- 矩估计: $p=m/n$. (MVU 估计)
- 极大似然估计: $p=m/n$. (MVU 估计)
- 贝叶斯估计
  - 同等无知原则: $p=\dfrac{X+1}{n+2}$.
  - 若先验密度 $h(p)=p^{a-1}(1-p)^{b-a-1}$, 则 $\tilde{p}=\dfrac{X+c}{n+d}$.
- 区间估计
  - 大样本法: 近似取枢轴变量 $\dfrac{X-np}{\sqrt{np(1-p)}}\sim N(0,1)$, 解不等式得

$$\left(1+\frac{z_{\alpha/2}^2}{n}\right)p^2-\left(2X+\frac{z_{\alpha/2}^2}{n}\right)p+X^2<0$$

$$\begin{aligned}\hat{p}_1,\hat{p}_2&=\frac{n}{n+z_{\alpha/2}^2}\left(\overline{X}+\frac{z_{\alpha/2}^2}{2n}\pm z_{\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}+\frac{z_{\alpha/2}^2}{4n^2}}\right)\\&\approx\overline{X}\pm z_{\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}.\end{aligned}$$

- $p^k\ (k\le n)$ 的无偏估计是 $\dfrac{X^{\underline{k}}}{n^{\underline{k}}}$. (下降阶乘幂)

2 泊松分布

2.1 基础概念

- $X\sim P(\lambda)$.
- 理解: 单位时间内事件平均发生 λ 次, 则某一段单位时间内发生的次数为 x.
- 概率分布: $P(X=i)=\lim\limits_{n\to\infty} b(i;n,\dfrac{\lambda}{n})=\dfrac{\mathrm{e}^{-\lambda}\lambda^i}{i!}$.
- 当二项分布满足 $n>50,p<0.1,np<5$ 时, 用泊松分布近似效果较好.

2.2 数字特征

- 最可能数: $k=\lfloor\lambda\rfloor$.
- 期望: $E(X)=\lambda$.
- 方差: $\mathrm{Var}(X)=\lambda$.
- 中位数: $m_e=\dfrac{\ln 2\lambda}{\lambda}$.
- $E\left|X-m_e\right|=m_e$.
- 母函数: $G(s)=\mathrm{e}^{\lambda(s-1)},\ s\in(-\infty,+\infty)$.
- 特征函数: $g(t)=\mathrm{e}^{\lambda\left(\mathrm{e}^{\mathrm{i}t}-1\right)}$.

2.3 其它性质

- 泊松分布和的函数 (可加性)

  $$X_1\sim P(\lambda_1),\ X_2\sim P(\lambda_2)\quad\Rightarrow\quad X_1+X_2\sim P(\lambda_1+\lambda_2).$$

- 记 $f(\lambda)=P(X\le k)$, 则 $f'(\lambda)<0$, 并且 $f(\lambda)=\dfrac{1}{k!}\displaystyle\int_\lambda^{+\infty} t^k\mathrm{e}^{-t}\,\mathrm{d}t$.

  注: 上式可用于参数检验.

- $P_\lambda(X \le k) = \sum_{i=0}^{k} \frac{\mathrm{e}^{-\lambda}\lambda^i}{i!} = \int_\lambda^{+\infty} \frac{\mathrm{e}^{-t}t^k}{k!}\,\mathrm{d}t = K_{2k+2}(2\lambda)$. (卡方分布函数)
- 若 $X \sim P(\lambda)$, $Y \sim B(X, p)$, 则 $Y \sim P(\lambda p)$.
- 若有一批零件寿命服从指数分布, 固定一个时间 $T > 0$, 让一个元件从时刻 0 开始工作, 每当这个元件坏了的时候马上用一个新的替换, 则到 T 时替换的次数 $X \sim P(\lambda T)$, 即 $P(X = n) = \frac{\mathrm{e}^{-\lambda T}(\lambda T)^n}{n!}$.
- 泊松分布的一个应用见特殊函数笔记中的 Dobinski 公式.

2.4 参数估计

- 矩估计
  - $\lambda = m$. (MVU 估计)
  - $\lambda = m_2$ 或 S^2.
- 极大似然估计: $\lambda = \overline{X}$.
- 贝叶斯估计: 见第四章第五题.
- 区间估计
  - 大样本法: 近似地取 $(Y_n - n\lambda)/\sqrt{n\lambda} \sim N(0, 1)$, 则

$$A, B = \overline{X} + u_{\alpha/2}^2/(2n) \pm u_{\alpha/2}\sqrt{u_{\alpha/2}^2/(4n^2) + \overline{X}/n}, \quad \overline{X} = Y_n/n.$$

3 超几何分布

3.1 基础概念

- $X \sim H(N, n, M)$.
- 理解: N 件产品中有 M 件次品, 从总体中抽 n 件时次品的数量 m.
- 概率分布: $P(X = m) = \binom{M}{m}\binom{N-M}{n-m}\Big/\binom{N}{n}$.

3.2 数字特征

- 期望: $E(X) = \frac{nM}{N}$.
- 方差: $\mathrm{Var}(X) = \frac{nM(N-n)(N-M)}{N^2(N-1)} = \frac{nM}{N}\frac{N-n}{N-1}\left(1 - \frac{M}{N}\right)$.

3.3 其它性质

3.4 参数估计

已知 N, n 估计 M.

- 贝叶斯估计: 采用同等无知原则, 则$M = \frac{N+2}{n+2}(X+1) - 1$.

4 负二项分布

4.1 基础概念

- $X \sim NB(r, p)$, 又称为正整数形式帕斯卡分布.
- 理解: 合格率为 p, 抽取到 r 个合格产品时, 抽到的不合格产品的个数 x.
- 概率分布: $P(X = i) = d(i; r, p) = \binom{i+r-1}{r-1}p^r(1-p)^i$.

4.2 数字特征

- 数学期望: $E(X)=\dfrac{r(1-p)}{p}$.
- 方差: $\operatorname{Var}(X)=\dfrac{r(1-p)}{p^2}$.

4.3 其它性质

4.4 参数估计

注: $m_e:=(X_1+X_2+\cdots+X_n)/n$.

- 矩估计: $p=\dfrac{r}{m_e+r}$.
- 极大似然估计: $p=\dfrac{r}{m_e+r}$.
- 贝叶斯估计: $p=\dfrac{nr+1}{nr+nm_e+1}$.

5 几何分布

5.1 基础概念

- $X\sim GE(p)$.
- 理解: 合格率为 p, 抽取到第一个合格产品时, 抽到的不合格产品的个数 x.
- 概率分布: $P(X=i)=p(1-p)^i$.
- 累积分布函数: $P(X\le k)=1-(1-p)^{k+1}$.
- 互补累积分布函数: $P(X\ge k)=(1-p)^k$.

5.2 数字特征

- 数学期望: $E(X)=\dfrac{1-p}{p}$.
- 方差: $\operatorname{Var}(X)=\dfrac{1-p}{p^2}$.
- 母函数: $G(s)=\dfrac{ps}{1-qs}-1,\ s\in\left(-\dfrac{1}{q},\dfrac{1}{q}\right)$.
- 特征函数: $g(t)=\dfrac{pe^{it}}{1-qe^{it}}-1$.

5.3 其它性质

- 几何分布具有无记忆性.
- 若 $X_1,X_2,\cdots,X_r$ 独立同分布 $GE(p)$, 则 $X_1+X_2+\cdots+X_r\sim NB(r,p)$.
- 若 $X_1\sim \mathrm{GE}(1-p_1)$ 和 $X_2\sim \mathrm{GE}(1-p_2)$ 独立, 则

$$\begin{aligned}&\min(X_1,X_2)\sim \mathrm{GE}(1-p_1p_2)\\&\max(X_1,X_2)\sim P(X=k)=p_1^k(1-p_1)+p_2^k(1-p_2)+p_1^kp_2^k(p_1p_2-1)\end{aligned}$$

  更一般的, 若 $X_i\sim \mathrm{GE}(1-p_i)$, 则 $\min_i(X_i)=\mathrm{GE}(1-\prod_i p_i)$.

5' 几何分布

5'.1 基础概念

- $X\sim G(p)$.
- 理解: 合格率为 p, 抽取到第一个合格产品时, 抽取的总产品的个数 x.
- 概率分布: $P(X=i)=p(1-p)^{i-1}$.

5'.2 数字特征

- 数学期望: $E(X)=\dfrac{1}{p}$.
- 方差: $\mathrm{Var}(X)=\dfrac{1-p}{p^2}$.
- 母函数
  - $G(s)=\dfrac{ps}{1-qs}$, $s\in\left(-\dfrac{1}{q},\dfrac{1}{q}\right)$.
  - $G^{(n)}(1)=\dfrac{(1-p)^{n-1}}{p^n}n!$.
- 特征函数: $g(t)=\dfrac{p\mathrm{e}^{\mathrm{i}t}}{1-q\mathrm{e}^{\mathrm{i}t}}$.

5'.3 其它性质

- 几何分布具有无记忆性.
- 若 $X_1,X_2,\cdots,X_r$ 独立同分布 $G(p)$, 则 $X_1+X_2+\cdots+X_r-r\sim NB(r,p)$.

6 泽塔分布

6.1 基础概念

- $X\sim\mathrm{Zeta}(s)$.
- Riemann Zeta 函数: $\zeta(s)=\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^s}$.
- 概率密度函数: $P(X=k)=\dfrac{1}{\zeta(s)k^s}$, $k=1,2,\cdots$.

6.2 数字特征

- k 阶矩: $E(X^k)=\dfrac{\zeta(s-k)}{\zeta(s)}$, $s>k+1$.
- 对数期望: $E(\ln X)=-\dfrac{\zeta'(s)}{\zeta(s)}$, $s>1$.
- 信息熵: $H(X)=E(-\ln(\mathrm{Zeta}(s)))=-\displaystyle\sum_{k=1}^{\infty}\frac{\ln\frac{1}{\zeta(s)k^s}}{\zeta(s)k^s}=\ln\zeta(s)-s\frac{\zeta'(s)}{\zeta(s)}$.

6.3 其它性质

- **问题 1 (最大熵分布)**

  对于取值为正整数的概率分布, 求给定对数期望的条件下熵最大的分布, 即

$$\max=-\sum_k p_k\ln p_k$$

$$\begin{cases}\displaystyle\sum_k p_k=1,\\ \displaystyle\sum_k p_k\ln k=a.\end{cases}$$

  由 Lagrange 乘数法解得此分布即为 Zeta 分布.

- **性质 1**

  设 $X\sim\mathrm{Zeta}(s)$, 则素因数分解中素数 p 的指数满足:

$$\nu_p(X)\sim\mathrm{GE}(1-p^{-s}).$$

  证明

$$
\begin{aligned}
P(\nu_p(X) \geq k) &= \frac{1}{\zeta(s)} \sum_{n=1}^{\infty} \frac{1}{(p^k n)^s} = \frac{1}{p^{ks}} \\
P(\nu_p(X) = k) &= \frac{1}{p^{ks}} - \frac{1}{p^{(k+1)s}} = (1-p^{-s})p^{-ks}
\end{aligned}
$$

- **性质 2**

  设 $X \sim \mathrm{Zeta}(s)$, 若 p 和 q 是两个互素的素数, 则 $\nu_p(X)$ 和 $\nu_q(X)$ 独立.

  证明

$$
P(\nu_p(X) \geq k, \nu_q(X) \geq l) = \frac{1}{(p^k q^l)^s} = \frac{1}{p^{ks} p^{ls}} = P(\nu_p(X) \geq k) P(\nu_q(X) \geq l).
$$

- **性质 3**

  设 $\mathbb{P}$ 是全体素数的集合, $\{X_p\}_{p \in \mathbb{P}}$ 是一组相互独立的随机变量, 其中 $X_p \sim \mathrm{GE}(1-p^{-s})$, 则

$$
Z = \prod_{p \in \mathbb{P}} p^{X(1-p^{-s})} \sim \mathrm{Z}(s).
$$

  证明 由 Euler 乘积公式 $\dfrac{1}{\zeta(s)} = \prod_{p \in \mathbb{P}} \left(1 - \dfrac{1}{p^s}\right)$ 得:

$$
P(Z = k) = \prod_{p \in \mathbb{P}} \left(1 - \frac{1}{p^s}\right) \frac{1}{p^{\nu_p(k)s}} = \frac{1}{\zeta(s)} \prod_{p \in \mathbb{P}} \frac{1}{p^{\nu_p(x)s}} = \frac{1}{\zeta(s) k^s}.
$$

- **性质 4**

  若 $X_1 \sim \mathrm{Zeta}(s_1)$ 和 $X_2 \sim \mathrm{Zeta}(s_2)$ 独立, 则

$$
\gcd(X_1, X_2) \sim \mathrm{Zeta}(s_1 + s_2).
$$

  证明

$$
\nu_p\left(\gcd(X_1, X_2)\right) = \min\left\{\nu_p(X_1), \nu_p(X_2)\right\} \sim \mathrm{GE}\left(1 - p^{-(s_1+s_2)}\right).
$$

- **问题 2 (两个随机的正整数互素的概率)**

  X 和 Y 互素 $\quad \Leftrightarrow \quad \gcd(X, Y) = 1.$

  正整数集上的均匀分布 $\sim \lim\limits_{s \to 1^+} \mathrm{Zeta}(s)$.

$$
\lim_{s_1, s_2 \to 1^+} P(\gcd(X, Y) = 1) = \lim_{s_1, s_2 \to 1^+} \frac{1}{\zeta(s_1 + s_2)} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}.
$$

  注: 这并非严格的证明.

A.2.2 一维连续型

1 正态分布

点击查看 Geogebra 图像 　　或直接打开 [网页链接]

1.1 基础概念

- 正态分布又称高斯分布.
- $X\sim N(\mu,\sigma^2)$.
- 概率密度函数: $f(x)=(\sqrt{2\pi}\sigma)^{-1}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$.
- 标准正态分布: $Y=(X-\mu)/\sigma\sim N(0,1)$.
- 3σ 原则: 0.6826，09544，9.9974.
- 上 α 分位数: $\Phi(z_\alpha)=1-\alpha$.

1.2 数字特征

- 期望: μ.
- 方差: σ^2.
- 二阶原点矩: $\alpha_2=E(X^2)=\sigma^2+\mu^2$.
- k 阶中心矩: $\mu_k=\begin{cases}\sigma^k(k-1)!!, & k\text{ 为偶数},\\ 0, & k\text{ 为奇数}.\end{cases}$
- 偏度系数: $\beta_1=0$.
- 峰度系数: $\beta_2=3$.
- 特征函数: $g(t)=\mathrm{e}^{\mathrm{i}\mu t-\frac{\sigma^2}{2}t^2}$.
- 矩母函数: $\phi_X(t)=\mathrm{e}^{\frac{t^2\sigma^2}{2n}+\mu t}$.

1.3 其它性质

- $aN(\mu,\sigma^2)+b\sim N(a\mu+b,a^2\sigma^2)$.
- 若 X 和 Y 独立同分布 $N(0,1)$, 则将 (X,Y) 化为极坐标 (R,Θ) 后, R 与 Θ 独立.
- 相互独立的正态分布的函数
  - 分布之和
    - 若 $X_1\sim N(\mu_1,\sigma_1^2),\ X_2\sim N(\mu_2,\sigma_2^2)$ 相互独立, 则 $X_1+X_2\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.
    - 若 $X_i\sim N(\mu_i,\sigma_i^2)$ 且相互独立, 则 $X_1+\cdots+X_n\sim N(\mu_1+\cdots+\mu_n,\sigma_1^2+\cdots+\sigma_n^2)$.
  - 分布之差

    若 $X_1\sim N(\mu_1,\sigma_1^2),\ X_2\sim N(\mu_2,\sigma_2^2)$ 相互独立, 则 $X_1-X_2\sim N(\mu_1-\mu_2,\sigma_1^2+\sigma_2^2)$.
  - 分布之商

    若 X_1 和 X_2 独立同分布 $N(0,1)$, 则 $X_1/X_2\sim C(1,0)$ (柯西分布).
  - 分布之积

    若 $X_1\sim N(0,\sigma_1^2),\ X_2\sim N(0,\sigma_2^2)$, 则 $X_1X_2\sim\dfrac{1}{\pi\sigma_1\sigma_2}K_0\left(\dfrac{|z|}{\sigma_1\sigma_2}\right)$ (修正贝塞尔函数; 暂时未学)
  - 平方之和

    若 $X_1,X_2,\cdots,X_n$ 独立同分布 $N(0,1)$, 则 $Y=X_1^2+X_2^2+\cdots+X_n^2\sim\chi_n^2$.
- 统计量的分布
  - $\overline{X}$ 与 $S^2=\dfrac{1}{n-1}\displaystyle\sum_{i=1}^n(X_i-\overline{X})$ 独立.
  - 均值已知, 标准差已知
    - $\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$.
    - $\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\sim N(0,1)$.
    - $\displaystyle\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi_n^2$.
  - 均值已知, 标准差未知

- $\dfrac{\sqrt{n}\,(\overline{X}-\mu)}{S} \sim t_{n-1}.$
- 均值未知, 标准差已知
  - $\dfrac{\mathrm{SS}}{\sigma^2} = \dfrac{(n-1)S^2}{\sigma^2} = \sum\limits_{i=1}^{n}\left(\dfrac{X_i-\overline{X}}{\sigma}\right)^2 \sim \chi^2_{n-1}.$
- 两份相互独立的样本

  $X_1, X_2, \cdots, X_{n_1}$, iid, $\sim N(\mu_1, \sigma_1^2)$.

  $Y_1, Y_2, \cdots, Y_{n_2}$, iid, $\sim N(\mu_2, \sigma_2^2)$.
  - $\overline{X}-\overline{Y} \sim N\left(\mu_1-\mu_2, \dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}\right).$
  - $\dfrac{S_1^2}{\sigma_1^2}\Big/\dfrac{S_2^2}{\sigma_2^2} \sim f(n_1-1, n_2-1).$
  - 当 $\sigma_1^2=\sigma_2^2=\sigma^2$ 时,

$$S_\omega := \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}},$$
$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_\omega\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t_{n_1+n_2-2}.$$

    注: 利用 $\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma}$ 和 $\dfrac{\mathrm{SS}}{\sigma^2}$, 由 t 分布的定义即得.

1.4 参数估计

- 单个正态总体 $N(\mu,\sigma^2)$ 均值 μ 与方差 σ^2 的估计.
  - 已知 σ^2, 估计 μ.
    - 矩估计: $\mu = m$.

      注: 无论 σ^2 是否已知, 均为 MVU 估计.
    - 区间估计 (枢轴变量法)

      根据 $\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \sim N(0,1)$, 知

$$[\hat{\mu}_1, \hat{\mu}_2] = \left[\overline{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}, \overline{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right].$$

  - 已知 μ, 估计 σ^2.
    - 矩估计: $\hat{\sigma^2} = m_2$.

      注: 这是 μ 已知时的 MVU 估计, 且此时均方误差为 $\dfrac{2}{n}\sigma^4$.
    - 区间估计 (枢轴变量法)

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

$$[\hat{\sigma_1^2}, \hat{\sigma_2^2}] = \left[\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_n\left(\dfrac{\alpha}{2}\right)}, \frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_n\left(\dfrac{\alpha}{2}\right)}\right].$$

  - 估计 μ 和 σ^2.
    - 矩估计

        - $\mu = m$.

          注: 无论 σ^2 是否已知, 均为 MVU 估计.
        - $\sigma^2 = S^2$.

          注: 这是 μ 未知时的 MVU 估计.
    - 极大似然估计
        - $\mu = m$. (MVU 估计)
        - $\sigma^2 = m_2$. (非 MVU 估计)
    - 区间估计 (枢轴变量法)
        - 根据 $\dfrac{\sqrt{n}\left(\overline{X}-\mu\right)}{S} \sim t_{n-1}$, 知一样本 t 区间估计为

$$[\hat{\mu}_1, \hat{\mu}_2] = \left[\overline{X} - \frac{S}{\sqrt{n}} t_{n-1}\left(\frac{\alpha}{2}\right), \overline{X} + \frac{S}{\sqrt{n}} t_{n-1}\left(\frac{\alpha}{2}\right)\right].$$

        - 根据 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$, 知

$$[\hat{\sigma_1^2}, \hat{\sigma_2^2}] = \left[\frac{(n-1)S^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)S^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)}\right].$$

    - 无偏估计 (通过调整系数而得)

$$\tilde{\sigma} = \sqrt{\frac{n-1}{2}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} S.$$

- 两个正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的均值差 $\mu_1 - \mu_2$ 与方差比 σ_1^2/σ_2^2 的区间估计.
    - 估计 $\delta = \mu_1 - \mu_2$.
        - 方差 σ_1^2 和 σ_2^2 已知.

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

$$[\hat{\delta}_1, \hat{\delta}_2] = \left[\overline{X} - \overline{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \right.$$
$$\left. \overline{X} - \overline{Y} + z_{\alpha/2} + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right].$$

        - 方差 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知.

$$S_\omega^2 := \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2},$$

$$\frac{(\overline{X}-\overline{Y}) - (\mu_1-\mu_2)}{S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2},$$

$$[\hat{\delta}_1, \hat{\delta}_2] = \left[\overline{X} - \overline{Y} - t_{n_1+n_2-2}\left(\frac{\alpha}{2}\right) S_\omega \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}, \right.$$
$$\left. \overline{X} - \overline{Y} + t_{n_1+n_2-2}\left(\frac{\alpha}{2}\right) S_\omega \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right].$$

- 方差 σ_1^2 和 σ_2^2 未知.

  即贝伦斯 - 费歇尔问题, 目前还没有较好的处理方法.

  不过可以利用大样本法, 近似同方差已知的情况处理.

$$
\begin{aligned}
N(0,1) &\sim \left[(\overline{X}-\overline{Y})-(\mu_1-\mu_2)\right]\Big/\sqrt{\sigma_1^2/n+\sigma_2^2/m} \qquad &(\text{严格的}) \\
&\sim \left[(\overline{X}-\overline{Y})-(\mu_1-\mu_2)\right]\Big/\sqrt{S_1^2/n+S_2^2/m} \qquad &(\text{近似的})
\end{aligned}
$$

- 估计 $\lambda=\sigma_1^2/\sigma_2^2$.
  - 均值 μ_1 和 μ_2 已知.

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

$$
[\hat{\lambda}_1,\hat{\lambda}_2]=\left[\frac{\sum\limits_{i=1}^{n_1}\dfrac{(X_i-\mu_1)^2}{n_1}}{\sum\limits_{i=1}^{n_2}\dfrac{(Y_i-\mu_2)^2}{n_2}}F_{n_2,n_1}\left(1-\frac{\alpha}{2}\right),\right.
\left.\frac{\sum\limits_{i=1}^{n_1}\dfrac{(X_i-\mu_1)^2}{n_1}}{\sum\limits_{i=1}^{n_2}\dfrac{(Y_i-\mu_2)^2}{n_2}}F_{n_2,n_1}\left(\frac{\alpha}{2}\right)\right].
$$

  - 均值 μ_1 和 μ_2 未知.

$$
\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F_{n_1-1,n_2-1},
$$

$$
[\hat{\lambda}_1,\hat{\lambda}_2]=\left[\frac{S_1^2}{S_2^2}F_{n_2-1,n_1-1}\left(1-\alpha\right),\frac{S_1^2}{S_2^2}F_{n_2-1,n_1-1}\left(\frac{\alpha}{2}\right)\right].
$$

- 估计变异系数 σ/μ.
  - 矩估计: $\sqrt{m_2}/m$ 或 S/m.
- 估计 $N(\theta,1)$ 的 θ.
  - 贝叶斯估计: 先验密度 $h(\theta)\sim N(\mu,\sigma^2)$, 则

    $$\tilde{\theta}=\frac{n}{n+1/\sigma^2}\overline{X}+\frac{1/\sigma^2}{n+1\sigma^2}\mu.$$

2 指数分布

点击查看 Geogebra 图像　　或直接打开 网页链接

2.1 基础概念

- 指数分布又称为负指数分布.
- $X\sim E(\lambda)$.
- 概率密度函数: $f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\le 0.\end{cases}$
- 分布函数: $F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\le 0.\end{cases}$

2.2 数字特征

- 数学期望: $E[X] = \lambda^{-1}$.
- 方差: $\mathrm{Var}[X] = \lambda^{-2}$.
- k 阶原点矩: $E(X^k) = \dfrac{k!}{\lambda^k}$.
- 特征函数: $g(t) = \dfrac{\lambda}{\lambda - \mathrm{i}t}$.
- 矩量母函数: $m_X(t) = \dfrac{\lambda}{\lambda - t}$, $t < \lambda$.

2.3 其它性质

- $aE(\lambda) = E\left(\dfrac{\lambda}{a}\right)$.
- 指数分布具有无记忆性，即 $P(X > m + t \mid X > m) = P(X > t)$.
- 若有一批元件寿命 $X \sim E(\lambda)$, 让一个元件开始工作, 每当这个元件坏了就用一个新的替换, 则到经历时间 T 后替换的次数 $Y \sim P(\lambda T)$.
- 若 $X_1, X_2, \cdots, X_n$ 独立同分布 $E(\lambda)$, 则

$$Y = 2\lambda(X_1 + X_2 + \cdots + X_n) \sim \chi^2_{2n}.$$

- 若 $X_i \sim E(\lambda_i)$ 相互独立, 则

$$Y = \min(X_1, X_2, \cdots, X_n) \sim E(\lambda_1 + \lambda_2 + \cdots + \lambda_n).$$

2.4 参数估计

- 矩估计: $1/\lambda = m$. (MVU 估计)
- 极大似然估计: $\lambda = 1/m$.
- 贝叶斯估计: 若先验密度为 $h(\lambda) = \lambda \mathrm{e}^{-\lambda}$ $(\lambda > 0)$, 其它值为零, 则 $\lambda = \dfrac{n+2}{n\overline{X}+1}$.
- 区间估计
  - 枢轴变量法
    - 估计 λ.

      由 $2n\lambda\overline{X} \sim \chi^2_{2n}$, 知

$$[\hat{\theta}_1, \hat{\theta}_2] = \left[\chi^2_{2n}(1-\alpha/2)/(2n\overline{X}),\ \chi^2_{2n}(\alpha/2)/(2n\overline{X})\right].$$

    - 估计 $1/\lambda$.

      由 $2n\lambda\overline{X} \sim \chi^2_{2n}$, 知

$$[\hat{\theta}_1, \hat{\theta}_2] = \left[(2n\overline{X})/\chi^2_{2n}(1-\alpha/2),\ (2n\overline{X})/\chi^2_{2n}(\alpha/2)\right].$$

- 若 $X_1, X_2, \cdots, X_n$ 独立同分布 $E(\lambda_1)$, $Y_1, Y_2, \cdots, Y_m$ 独立同分布 $E(\lambda_2)$, 估计 λ_2/λ_1.
  - 区间估计 (枢轴变量法)
    - 利用 $\dfrac{4\lambda_1 n\overline{X}}{4\lambda_2 m\overline{Y}} \sim \dfrac{2n\chi^2_{2n}}{2m\chi^2_{2m}} \sim F_{2n,2m}$.

3 混合指数分布

3.1 基础概念

- 混合指数分布又称为 **超指数分布** (Hyperexponential Distribution).
- 理解: 设有 m 个平行的服务台 $X_i \sim E(\lambda_i)$, 若顾客有 p_i 的概率选取第 i 个服务台, 则这样顾客的服务时间分布服从 m 阶超指数分布.
- 概率密度函数 $f(x)=\sum_{i=1}^{m} p_i \lambda_i \mathrm{e}^{-\lambda_i x}, \quad (x>0).$

  其中 $\sum_{i=1}^{m} p_i=1.$
- 累积分布函数 $F(x)=\sum_{i=1}^{m} p_i\left(1-\mathrm{e}^{-\lambda_i x}\right), \quad(x>0).$

3.2 数字特征

- 和指数分布一样, 不再赘述. 如

  k 阶原点矩 $E(X^k)=\sum_{i=1}^{m} \frac{k!}{\lambda_i^k} p_i.$

3.3 其它性质

- 无记忆性.
- 若 $p_i=\frac{1}{m}$, $i=1,2,\cdots,m$, 则 $X \sim \chi_{2m}^2$.

3.4 参数估计

- 矩估计.
- 优化估计.

4 均匀分布

4.1 基础概念

- $X \sim R(a, b)$.
- 概率密度函数: $f(x)=\begin{cases}1/(b-a), & a \leq x \leq b, \\ 0, & x<a \text{ 或 } x>b.\end{cases}$
- 分布函数: $F(x)=\begin{cases}0, & x \leq a, \\ (x-a)/(b-a), & a<x<b, \\ 1, & x \geq b.\end{cases}$

4.2 数字特征

- 数学期望: $\frac{a+b}{2}$.
- 方差: $\frac{(b-a)^2}{12}$.
- k 阶原点矩: $\alpha_k=\frac{1}{k+1} \frac{b^{k+1}-a^{k+1}}{b-a}$.
- k 阶中心距: $\mu_k=\begin{cases}\frac{1}{k+1}\left(\frac{b-a}{2}\right)^k, & k \text{ 为偶数}, \\ 0, & k \text{ 为奇数}.\end{cases}$
- 偏度系数: $\beta_1=0$.
- 峰度系数: $\beta_2=\frac{9}{5}$.
- 特征函数: $g(t)=\begin{cases}\frac{\mathrm{e}^{\mathrm{i} b t}-\mathrm{e}^{\mathrm{i} a t}}{\mathrm{i} t(b-a)}, & t \neq 0, \\ 1, & t=0.\end{cases}$

4.3 其它性质

- $cR(a,b)+d \sim R(ac+d, bc+d)\ (c>0)$.
- 若 $X_1, X_2, \cdots, X_n$ 独立同分布 $U(a,b)$, 则

$$\max(X_1+X_2+\cdots+X_n) \sim f(x) = \frac{n(x-a)^{n-1}}{(b-a)^n}$$
$$\min(X_1+X_2+\cdots+X_n) \sim f(x) = \frac{n(b-x)^{n-1}}{(b-a)^n}$$

- 若 $X \sim U\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, 则 $\tan X \sim C(1,0)$.

4.4 参数估计

- 估计 $R(\theta_1, \theta_2)$ 的参数.
  - 矩估计: $\theta_1 = m - \sqrt{3m_2}$, $\theta_2 = m + \sqrt{3m_2}$.
  - 极大似然估计: $\theta_1 = \min_i(X_i)$, $\theta_2 = \max_i(X_i)$.
- 估计 $R(0,\theta)$ 的参数.
  - 极大似然估计: $\hat{\theta} = \max_i(X_i)$.
  - 无偏估计
    - $\hat{\theta} = \frac{n+1}{n}\max_i(X_i)$. (MVU 估计, 也是相合估计)
    - $\hat{\theta} = (n+1)\min_i(X_i)$. (方差很大)
    - $\hat{\theta} = \max_i(X_i) + \min_i(X_i)$.
  - 区间估计
    - 由 $\hat{\theta}_1 := \max_i(X_i) \sim F_{\hat{\theta}_1}(x) = \frac{nx^{n-1}}{\theta^n}$,

      $[\max(X_i), (1-\alpha)^{-\frac{1}{n}}\max(X_i)]$ 的置信系数为 $1-\alpha$.

5 对数正态分布

点击查看 Geogebra 图像　　或直接打开 网页链接

5.1 基础概念

- $\ln X \sim N(\mu, \sigma^2)$.
- 概率密度函数: $f(x,\mu,\sigma) = \begin{cases} \left(x\sqrt{2\pi}\sigma\right)^{-1}\exp\left[-\frac{(\ln x-\mu)^2}{2\sigma^2}\right], & x>0, \\ 0, & x \le 0. \end{cases}$

5.2 数字特征

- 期望: $E(X) = \mathrm{e}^{\mu+\sigma^2/2}$.
- 方差: $\mathrm{Var}(X) = \left(\mathrm{e}^{\sigma^2}-1\right)\mathrm{e}^{2\mu+\sigma^2} = \left(\mathrm{e}^{\sigma^2}-1\right)E(X)^2$.
- k 阶原点矩: $\alpha_k = \mathrm{e}^{\mu k + k^2\sigma^2/2}$.
- 偏度系数: $\beta_1 = \frac{\mu_3}{\sigma^3} = \frac{\mathrm{e}^{2\sigma^2}-3\mathrm{e}^{\sigma^2}+1}{\left(\mathrm{e}^{2\sigma^2}-1\right)^{3/2}}$.
- 峰度系数: $\beta_2 = \frac{\mu_4}{\sigma^4} = \left(\mathrm{e}^{\sigma^2}-1\right)\left(\mathrm{e}^{4\sigma^2}+2\mathrm{e}^{3\sigma^2}+3\mathrm{e}^{2\sigma^2}-3\right) > 0$.

5.3 其它性质

- $\ln bX^a \sim N(a\mu + \ln b, a^2\sigma^2)$.
- 对数正态分布总是右偏的.
- 对数正态分布的期望和方差都是两个参数的增函数.

  而在正态分布中, 期望与 σ 无关, 方差与 μ 无关.
- $\lim\limits_{\sigma \to 0+} E(X) = \mathrm{e}^{\mu}$.

  $\lim\limits_{\sigma \to 0+} \mathrm{Var}(X) = 0$.

  当 $\mu = 0$ 时, $E(X^k) = E(X)^{k^2}$.

5.4 参数估计

- 矩估计
  - $\hat{\sigma^2} = \ln\left(1 + \dfrac{S^2}{\overline{X}^2}\right)$.
  - $\hat{\mu} = 2\ln\overline{X} - \dfrac{1}{2}\ln\left(\overline{X}^2 + S^2\right)$.

6 柯西分布

点击查看 Geogebra 图像 或直接打开 网页链接

6.1 基础概念

- $X \sim C(\gamma, x_0)$.
- 概率密度函数: $f(x; x_0, \gamma) = \dfrac{1}{\pi} \cdot \dfrac{\gamma}{(x - x_0)^2 + \gamma^2} \quad (-\infty < x < +\infty)$.
- 累积分布函数: $F(x; x_0, \gamma) = \dfrac{1}{\pi}\arctan\dfrac{x - x_0}{\gamma} + \dfrac{1}{2}$.
- 标准柯西分布: $C(1, 0) \sim t_1$.
- 广义柯西分布: $X_k \sim f_m(X_k \mid \sigma_X) = \dfrac{a_m}{1 + \left(\dfrac{X_k^2}{2\sigma_k^2}\right)^m} \ (a_m > 0.5)$.

6.2 数字特征

- 数学期望不存在. (仅 Cauchy 主值积分存在)
- 方差不存在.
- 高阶矩不存在.

6.3 其它性质

- 可加性: 若 X_i 独立同分布 $C(\gamma, x_0)$, 则 $X_1 + X_2 + \cdots + X_n \sim C(n\gamma, nx_0)$.
- 若 X_1 和 X_2 独立同分布 $N(0, 1)$, 则 $\dfrac{X_1}{X_2} \sim C(1, 0)$.
- 若 $X \sim U\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$, 则 $\tan X \sim C(1, 0)$.

6.4 参数估计

- 参数估计: 可使用样本中位数 $\tilde{m}$ 估计.

7 拉普拉斯分布

点击查看 Geogebra 图像　　或直接打开 网页链接

7.1 基础概念

- $X \sim \mathrm{La}(\mu, \lambda)$.
- 概率密度函数: $f(x) = \dfrac{1}{2\lambda}\mathrm{e}^{-\frac{|x-\mu|}{\lambda}}$.
- 累积分布函数: $F(x) = \begin{cases} \dfrac{1}{2}\mathrm{e}^{\frac{x-\mu}{\lambda}}, & x < \mu. \\ 1 - \dfrac{1}{2}\mathrm{e}^{\frac{\mu-x}{\lambda}}, & x \geq \mu. \end{cases}$
- 参数说明
    - μ 是位置参数.
    - γ 是尺度参数, 越小曲线越陡.
    - 当 $\mu = 0$ 时, 正半部分是指数分布 $E\left(\lambda^{-1}\right)$ 概率密度的一半.

7.2 数字特征

- 矩及相关量
    - k 阶中心距: $\mu_k = E[(X-\mu)^k] = \begin{cases} 0, & k \text{ 为奇数}, \\ k!\lambda^k, & k \text{ 为偶数}. \end{cases}$
    - 期望 $E(X) = \mu$.
    - 方差 $\mathrm{Var}(X) = 2\lambda^2$.
- 同指数分布 (注意与 k 阶中心距、期望和方差比较)
    - $E(|X-\mu|^k) = k!\lambda^k$.
    - $E(|X-\mu|) = \lambda$.
    - $\mathrm{Var}(|X-\mu|) = \lambda^2$.
- 一些系数
    - 偏度系数 $\beta_1 = \dfrac{\mu_3}{\sigma^3} = 0$.
    - 峰度系数 $\beta_2 = \dfrac{\mu_4}{\sigma^4} = 6$.
- 相关函数
    - 矩量母函数 $m(t) = \dfrac{\mathrm{e}^{\frac{\mu t}{\lambda}}}{1-\lambda^2 t^2}$.
    - 特征函数 $g(t) = E(\mathrm{e}^{\mathrm{i}tX}) = \dfrac{\mathrm{e}^{\frac{\mathrm{i}\mu t}{\lambda}}}{1+\lambda^2 t^2}$.

7.3 其它性质

- $a\mathrm{La}(\mu, \lambda) + b \sim \mathrm{La}(a\mu + b, a\lambda)$.

  $\mathrm{La}(\mu, \lambda) \sim \lambda\mathrm{La}(0, 1) + \mu$.
- 注意到 Laplace 分布与指数分布的关系, 可以立即得到如下平凡的结论
    - 对于 $X \sim \mathrm{La}(\alpha, \beta)$, 有 $\dfrac{2}{\beta}|X-\alpha| \sim \chi_2^2 \sim \Gamma(1, 2)$.
    - 若 X 和 Y 独立同分布于 $\mathrm{La}(\alpha, \beta)$, 则 $\dfrac{|X-\alpha|}{|Y-\alpha|} \sim F_{2,2}$.
    - 若 $X_{11}, X_{12}, X_{21}, X_{22}$ 独立同分布于 $N(0, 1)$, 则 $D = \begin{vmatrix} X & Y \\ Z & W \end{vmatrix} \sim \mathrm{La}(0, 2)$.
- 与稳健性的联系

  古典回归分析中, 用偏差平方和的大小作为标准, 这种回归不具有稳健性.

  而改成偏差的绝对值和作为标准, 却具有稳健性 (尽管求解更加困难).
- 标准 Laplace 分布

- 概率密度: $\mathrm{La}(x;0,1)=\dfrac{\mathrm{e}^{-|x|}}{2}$.
- 特征函数: $\phi_L(t)=\dfrac{1}{1+t^2}$.

7.4 参数估计

- 估计 μ.
  - 矩估计: $\hat{\mu}=\overline{X}$.
  - 极大似然估计: $\hat{\mu}=m_e$ (中位数).
- 估计 λ.
  - 矩估计: $\hat{\lambda}=\dfrac{\sqrt{2}}{2}S$ (标准差).
  - 类似矩估计的估计:

$$\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}|X_i-\mu|\approx\frac{1}{n}\sum_{i=1}^{n}|X_i-\hat{\mu}|.$$

8 卡方分布

点击查看 Geogebra 图像 或直接打开 网页链接

8.1 基础概念

- 自由度为 n 的皮尔逊卡方密度与卡方分布 $X\sim\chi_n^2$.
- 概率密度函数

$$k_n(x)=\begin{cases}\dfrac{\mathrm{e}^{-x/2}x^{(n-2)/2}}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}, & x>0,\\ 0, & x\leq 0.\end{cases}$$

- 例子 (以下 $x>0$)

$$k_1(x)=\frac{\mathrm{e}^{-\frac{x}{2}}}{\sqrt{2\pi x}}\qquad k_2(x)=\frac{1}{2}\mathrm{e}^{-\frac{x}{2}}$$
$$k_3(x)=\frac{\sqrt{x}\mathrm{e}^{-\frac{x}{2}}}{\sqrt{2\pi}}\qquad k_4(x)=\frac{x}{4}\mathrm{e}^{-\frac{x}{2}}$$

- 上 α 分位数 $\chi_\alpha^2(n)$.
- 由中心极限定理近似求值 $X\sim\chi_n^2$,

$$\frac{X-n}{\sqrt{2n}}\dot{\sim}N(0,1)\quad\Rightarrow\quad\frac{\chi_\alpha^2(n)-n}{\sqrt{2n}}\approx z_\alpha\quad\Rightarrow\quad\chi_\alpha^2(n)\approx n+z_\alpha\sqrt{2n}.$$

8.2 数字特征

- $E(X)=n$.
- $\mathrm{Var}(X)=2n$.

  注意到方差是均值的两倍，可以以此检验是否为卡方分布.
- $E(X^{-1})=\dfrac{1}{n-2}$.
- $E(X^k)=\dfrac{2^k\Gamma\left(\frac{n}{2}+k\right)}{\Gamma\left(\frac{n}{2}\right)}\quad\left(k>-\dfrac{n}{2}\right)$.

- 特征函数: $g(t)=\dfrac{1}{(1-2\mathrm{i}t)^{\frac{n}{2}}}$.
- 矩量母函数: $m_X(t)=\dfrac{1}{(1-2t)^{\frac{n}{2}}},\quad t<\dfrac{1}{2}$.

8.3 其它性质

- 若 $X\sim t_n$, 则 $X^2\sim F_{1,n}$.
- 若 $X_1,X_2,\cdots,X_n$ 独立同分布 $N(0,1)$, 则

$$Y=X_1^2+X_2^2+\cdots+X_n^2\sim\chi_n^2.$$

- 若 $X_1\sim\chi_m^2$ 与 $X_2\sim\chi_n^2$ 独立, 则

$$X_1+X_2\sim\chi_{m+n}^2.$$

- 若 $X_1,X_2,\cdots,X_n$ 独立同分布 $E(\lambda)$, 则

$$X=2\lambda(X_1+X_2+\cdots+X_n)\sim\chi_{2n}^2.$$

- 若 $X_1,X_2,\cdots,X_n$ 独立同分布 $N(\mu,\sigma^2)$, 则

$$\frac{\mathrm{SS}}{\sigma^2}=\frac{(n-1)S^2}{\sigma^2}\sim\chi_{n-1}^2.$$

9 t 分布

点击查看 Geogebra 图像　　或直接打开 网页链接

9.1 基础概念

- 自由度为 n 的 t 分布 $X\sim t_n$.
- 概率密度函数

$$t_n(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}=\frac{\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}}{\mathrm{B}\left(\frac{n}{2},\frac{1}{2}\right)\sqrt{n}},\qquad -\infty<x<+\infty.$$

- 上 α 分位数 $t_\alpha(n)\approx z_\alpha$. (正态分布的上 α 分位数)
- 由对称性知: $t_{1-\alpha}(n)=-t_\alpha(n)$.

9.2 数字特征

- $E(t_n)=0\ (n>1)$.
- $\mathrm{Var}(t_n)=\dfrac{n}{n-2}\ (n>2)$.
- $E(X^k)=\dfrac{\mathrm{B}\left(\frac{n-k}{2},\frac{k+1}{2}\right)}{\mathrm{B}\left(\frac{n}{2},\frac{1}{2}\right)}n^{\frac{k}{2}}\ (-1<k<n)$.

9.3 其它性质

- 若 $X\sim N(0,1)$ 与 $Y\sim\chi_n^2$ 独立, 则

$$\frac{X}{\sqrt{Y/n}}\sim t_n.$$

- 若 $X_1,X_2,\cdots,X_n$ 独立同分布 $N(\mu,\sigma^2)$, 则

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}.$$

- 设 $X_1, X_2, \cdots, X_n$ 独立同分布 $N(\mu_1, \sigma^2)$, $Y_1, Y_2, \cdots, Y_m$ 独立同分布 $N(\mu_2, \sigma^2)$, 且 X_i, Y_j 独立, 则

$$\frac{\sqrt{\dfrac{nm(n+m-2)}{n+m}}\left[(\overline{X}+\overline{Y})-(\mu_1-\mu_2)\right]}{\sqrt{\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{j=1}^{m}(Y_j-\overline{Y})^2}} \sim t_{n+m-2}.$$

10 F 分布

点击查看 Geogebra 图像　　或直接打开 网页链接

10.1 基础概念

- 自由度为 (m,n) 的 F 分布 $X \sim F_{m,n}$.
- 概率密度函数

$$\begin{aligned} f_{m,n}(x) &= \mathrm{B}\left(\frac{m}{2}, \frac{n}{2}\right) m^{\frac{m}{2}} n^{\frac{n}{2}} x^{\frac{m}{2}-1}(mx+n)^{-\frac{m+n}{2}} \qquad (x>0) \\ &= \mathrm{B}\left(\frac{m}{2}, \frac{n}{2}\right)\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1}\left(1+\frac{m}{n}x\right)^{-\frac{m+n}{2}} \qquad (x>0) \end{aligned}$$

- 若 $F \sim F(m,n)$, 则 $F^{-1} \sim F(n,m)$.
- 第一自由度为 n_1, 第二自由度为 n_2 的 F 分布的上 α 分位数 $F_\alpha(n_1,n_2)$.
- $F_\alpha(m,n)\cdot F_{1-\alpha}(n,m)=1$.

10.2 数字特征

- $E(f_{m,n}) = \dfrac{n}{n-2}\ (n>2)$.
- $\mathrm{Var}(f_{m,n}) = \dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$.
- $E(X^k) = \left(\dfrac{n}{m}\right)^k \mathrm{B}\left(\dfrac{m}{2}, \dfrac{n}{2}\right)\mathrm{B}\left(\dfrac{m}{2}+k, \dfrac{n}{2}-k\right)$.

10.3 其它性质

- 设 X_1, X_2 独立, $X_1 \sim \chi^2_m$, $X_2 \sim \chi^2_n$, 则

$$\frac{X_1}{m}\bigg/\frac{X_2}{n} \sim F_{m,n}.$$

- 设 $X_1, X_2, \cdots, X_n$ 独立同分布 $N(\mu_1, \sigma_1^2)$, $Y_1, Y_2, \cdots, Y_m$ 独立同分布 $N(\mu_2, \sigma_2^2)$, 且 X_i, Y_j 独立, 则

$$\frac{S_Y}{\sigma_2^2}\bigg/\frac{S_X}{\sigma_1^2} \sim F_{m-1,n-1}.$$

- $\forall k, n, \in \mathbb{N}, a \in (0,1): kF_{k,n}(a) \geq F_{1,n}(a)$.

11 贝塔分布

点击查看 Geogebra 图像　　或直接打开 网页链接

11.1 基础概念

- $X \sim \mathrm{Be}(\alpha, \beta) \quad (\alpha, \beta > 0)$.
- 概率密度函数 $f(x; \alpha, \beta) = \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha, \beta)} \quad (0 < x < 1)$.
- 累积分布函数 $F(x; \alpha, \beta) = \dfrac{\mathrm{B}_x(\alpha, \beta)}{\mathrm{B}(\alpha, \beta)} = I_x(\alpha, \beta)$.
  - 不完全 Beta 函数 $\mathrm{B}_x(\alpha, \beta)$.
  - 正则不完全 Beta 函数 $I_x(\alpha, \beta)$.

11.2 数字特征

- 常用统计量
  - 众数 $M_0 = \dfrac{\alpha - 1}{\alpha + \beta - 2}$. (伯努利分布参数的极大似然估计)
  - 期望 $E(X) = \dfrac{\alpha}{\alpha + \beta}$. (伯努利分布参数的贝叶斯估计 & 同等无知原则)
  - 方差 $\mathrm{Var}(X) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.
- 矩及相关量
  - k 阶矩 $E(X^k) = \dfrac{\mathrm{B}(\alpha+k, \beta)}{\mathrm{B}(\alpha, \beta)} = \dfrac{(\alpha)^{(k)}}{(\alpha+\beta)^{(k)}} = \dfrac{\alpha+k-1}{\alpha+\beta+k-1} E(X^{k-1})$.
  - 偏度 $\beta_1 = \dfrac{\mu_3}{\sigma^3} = \dfrac{2(\beta-\alpha)\sqrt{\alpha+\beta+1}}{(\alpha+\beta+2)\sqrt{\alpha\beta}}$.
  - 峰度 $\beta_2 = \dfrac{\mu_4}{\mu_2^2} - 3 = \dfrac{6[(\alpha-\beta)^2(\alpha+\beta+1) - \alpha\beta(\alpha+\beta+2)]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)}$.

11.3 其它性质

- Beta 分布即伯努利分布的共轭先验分布.
- $E(\ln X) = \psi(\alpha) - \psi(\alpha + \beta)$.
  - $\mathrm{B}_{p,q}(\alpha, \beta) := \dfrac{\partial^{p+q}\mathrm{B}(x, y)}{\partial^p x \partial^q y} = \displaystyle\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} \ln^p x \ln^q (1-x) \, \mathrm{d}x$.
  - $\mathrm{B}_{1,0}(x, y) = \mathrm{B}(x, y)(\psi(x) - \psi(x+y))$.
  - $\mathrm{B}_{0,1}(x, y) = \mathrm{B}(x, y)(\psi(y) - \psi(x+y))$.
- Beta 分布与 Gamma 分布的关系.

12 伽马分布

点击查看 Geogebra 图像　　或直接打开 网页链接

12.1 基础概念

- $X \sim \mathrm{Ga}(\alpha, \beta) \sim \Gamma\left(\alpha, \dfrac{1}{\beta}\right)$.
- 概率密度函数 $f(x; \alpha, \beta) = \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\beta x} \quad (x > 0)$.
  - α 称为形状参数.
  - β 称为逆尺度参数.
- 累积分布函数 $F(x; \alpha, \beta) = \dfrac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}$.
  - 其中 $\gamma(s, x)$ 为下不完全 Gamma 函数.
  - 此外 $\Gamma(s, x)$ 为上不完全 Gamma 函数.
- 注意区分

- Gamma 分布 $\mathrm{Ga}(\alpha,\beta)$ 或 $\mathrm{Gamma}(\alpha,\beta)$, 其累积分布函数如上所示.
- Gamma 分布的另一种定义 $\Gamma(\alpha,\beta)$, 其累积分布函数为 $\dfrac{\gamma(\alpha,x/\beta)}{\Gamma(\alpha)}$.
- 上述定义的密度函数 $\Gamma(x;\alpha,\beta)$ 或 $\Gamma(X\mid\alpha,\beta)$.
- 上不完全 Gamma 函数 $\Gamma(s,x)$ 和 Gamma 函数 $\Gamma(s)$.

- 当 $\alpha\in\mathbb{N}$ 时, 退化为埃尔朗分布.

12.2 数字特征

- 有量纲参数
  - 众数 $M_0=\dfrac{\alpha-1}{\beta}\quad(\alpha>1)$.
  - k 阶原点矩 $\alpha_k=E(X^k)=\dfrac{\Gamma(\alpha+k)}{\beta^k\Gamma(\alpha)}=\dfrac{(\alpha)^{(k)}}{\beta^k}$.
  - 期望 $\mu=E(X)=\dfrac{\alpha}{\beta}$.
  - 方差 $\sigma^2=\mathrm{Var}(X)=\dfrac{\alpha}{\beta^2}$.
- 无量纲参数
  - 偏度系数 $\beta_1=\dfrac{\mu_3}{\sigma^3}=\dfrac{2}{\sqrt{\alpha}}$.
  - 峰度系数 $\beta_2=\dfrac{\mu_4}{\sigma^4}=\dfrac{6}{\alpha}$.
  - 变异系数 $c_v=\dfrac{\sigma}{\mu}=\dfrac{1}{\sqrt{\alpha}}$.
- 特征函数 $g(t)=\left(1-\dfrac{\mathrm{i}t}{\beta}\right)^{-\alpha}$.
- 矩量母函数 $m_X(t)=m_X(t)=\left(1-\dfrac{t}{\beta}\right)^{-\alpha}$, $t<\beta$.

12.3 其它性质

- 变化趋势
  - 当 $\alpha\in(0,1]$ 时, $f(x;\alpha,\beta)$ 递减.
  - 当 $\alpha\in(1,+\infty)$ 时, $f(x;\alpha,\beta)$ 先增后减, 为单峰函数.
  - 无量纲参数与图像形状仅与 α 有关, 故 α 称为形状参数.
- 特殊情况
  - 指数分布 $\mathrm{Ga}(1,\lambda)\sim E(\lambda)$.

    另一定义 $\Gamma\left(1,\dfrac{1}{\lambda}\right)\sim E(\lambda)$. (此外也有一种定义, 使得 $\Gamma(1,\lambda)\sim E(\lambda)$, 问就是别用)
  - 卡方分布 $\mathrm{Ga}\left(\dfrac{n}{2},\dfrac{1}{2}\right)\sim\chi^2_n$.

    另一定义 $\Gamma\left(\dfrac{n}{2},2\right)\sim\chi^2_n$.
- 函数运算
  - 数乘
    - 若 $X\sim\mathrm{Ga}(\alpha,\beta)$, 则 $\lambda X\sim\mathrm{Ga}\left(\alpha,\dfrac{\beta}{\lambda}\right)$.
    - 因此 β 称为尺度参数或逆尺度参数, 即 β 越大, 曲线越窄, 图像越接近 y 轴.
  - 可加性
    - 若 $X_1\sim\mathrm{Ga}(\alpha_1,\beta)$ 和 $X_2\sim\mathrm{Ga}(\alpha_2,\beta)$ 独立, 则 $X+Y\sim\mathrm{Ga}(\alpha_1+\alpha_2,\beta)$.
    - 特例 1 (卡方分布): 若 $X_m\sim\chi^2_m$, $X_n\sim\chi^2_n$ 独立, 则 $\chi^2_m+\chi^2_n\sim\chi^2_{m+n}$.
    - 特例 2 (正态分布): X_i 独立同分布 $N(0,1)$, 则 $X_1^2+X_2^2+\cdots+X_n^2\sim\chi^2_n$.
    - 特例 3 (指数分布): 若 X_i 独立同分布 $E(\lambda)$, 则 $2\lambda(X_1+X_2+\cdots+X_n)\sim\chi^2_{2n}$.

13 威布尔分布

点击查看 Geogebra 图像　　或直接打开 网页链接

13.1 基础概念

- 概率密度函数: $f(x;\lambda,k)=\begin{cases}\dfrac{k}{\lambda}\left(\dfrac{x}{\lambda}\right)^{k-1}\mathrm{e}^{-\left(\frac{x}{\lambda}\right)^k}, & x\geq 0,\\ 0, & x<0.\end{cases}$
- 累积分布函数: $F(x)=\begin{cases}1-\mathrm{e}^{-\left(\frac{x}{\lambda}\right)^k}, & x\geq 0,\\ 0, & x\leq 0.\end{cases}$

13.2 数字特征

- n 阶原点矩: $E(X^n)=\lambda^n\Gamma\left(1+\dfrac{n}{k}\right)$.
- 期望、方差、偏度、峰度等可由原点矩直接得到, 形式复杂故不再列出.

14 瑞利分布

14.1 基础概念

- 概率密度函数: $f(x)=\dfrac{x}{\sigma^2}\mathrm{e}^{-\frac{x^2}{2\sigma^2}},\quad x>0.$
- 累积分布函数: $F(x)=1-\mathrm{e}^{-\frac{x^2}{\lambda^2}},\quad x>0.$

14.2 数字特征

- k 阶原点矩: $E(X^k)=(2\sigma^2)^k\Gamma\left(1+\dfrac{k}{2}\right)$.
- 期望: $E(X)=\sqrt{\dfrac{\pi}{2}}\sigma\approx 1.253\sigma$.
- 方差: $\mathrm{Var}(X)=\dfrac{4-\pi}{2}\sigma^2\approx 0.429\sigma^2$.

15 帕累托分布

15.1 基础概念

- 帕累托分布 (Pareto Distribution) 又称布拉德福分布, 与幂律分布形式相同. 参考齐夫定律.
- 概率密度函数: $f(x)=\begin{cases}0, & x<x_{\min},\\ \dfrac{kx_{\min}^k}{x^{k+1}}, & x\geq x_{\min}.\end{cases}$
- 互补累积分布函数: $P(X>x)=\begin{cases}0, & x<x_{\min},\\ \left(\dfrac{x_{\min}}{x}\right)^k, & x\geq x_{\min}.\end{cases}$

  互补累积分布函数又称为*生存函数*, *残存函数* 或 *可靠性函数*.
- 大致服从帕累托分布的例子
  - 个人财富或资源的分布.
  - 人类居住区的大小.
  - 对百科条目的访问.
  - 龙卷风带来的灾难的数量.

15.2 数字特征

- 方便起见, 改变记号如下: $f(x)=\dfrac{\alpha\theta^\alpha}{x^{\alpha+1}},\ x\geq\theta.$
- k 阶矩: $E(X^k)=\dfrac{\alpha\theta^k}{\alpha-k},\ (\alpha>k).$
- 期望: $E(X)=\dfrac{\alpha\theta}{\alpha-1},\ (\alpha>1).$
- 方差: $\operatorname{Var}(X)=\dfrac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)},\ (\alpha>2).$

16 逻辑斯蒂分布

16.1 基础概念

- $X\sim L(\mu,\gamma).$
- Logistic 分布属于位置-尺度参数族.
- 累积分布函数: $F(x)=\dfrac{1}{1+\mathrm{e}^{-\frac{x-\mu}{\gamma}}}=\dfrac{1}{2}\left(1+\tanh\dfrac{x-\mu}{2\gamma}\right),\quad x\in\mathbb{R},\ \gamma>0.$
- 概率密度函数: $f(x)=\dfrac{\mathrm{e}^{-\frac{x-\mu}{\gamma}}}{\gamma\left(1+\mathrm{e}^{-\frac{x-\mu}{\gamma}}\right)^2}.$
- 参数说明
  - μ 是位置参数, 称为 **散布中心**.
  - γ 是尺度参数, 称为 **散布程度**.
  - 当 $\mu=0$ 时, $\gamma=\pm\gamma_0$ 的分布相同.
- 标准 Logistic 分布 $L(0,1)$.
  - 累积分布函数: $F_0(x)=\dfrac{1}{1+\mathrm{e}^{-x}}.$
  - 概率密度函数: $f_0(x)=\dfrac{\mathrm{e}^{-x}}{(1+\mathrm{e}^{-x})^2}.$

16.2 数字特征

- 期望 $E(X)=\mu.$
- 方差 $\operatorname{Var}(X)=\dfrac{\gamma^2\pi^2}{3}.$

16.3 其它性质

- $aL(\mu,\gamma)+b\sim L(a\mu+b,a\gamma).$

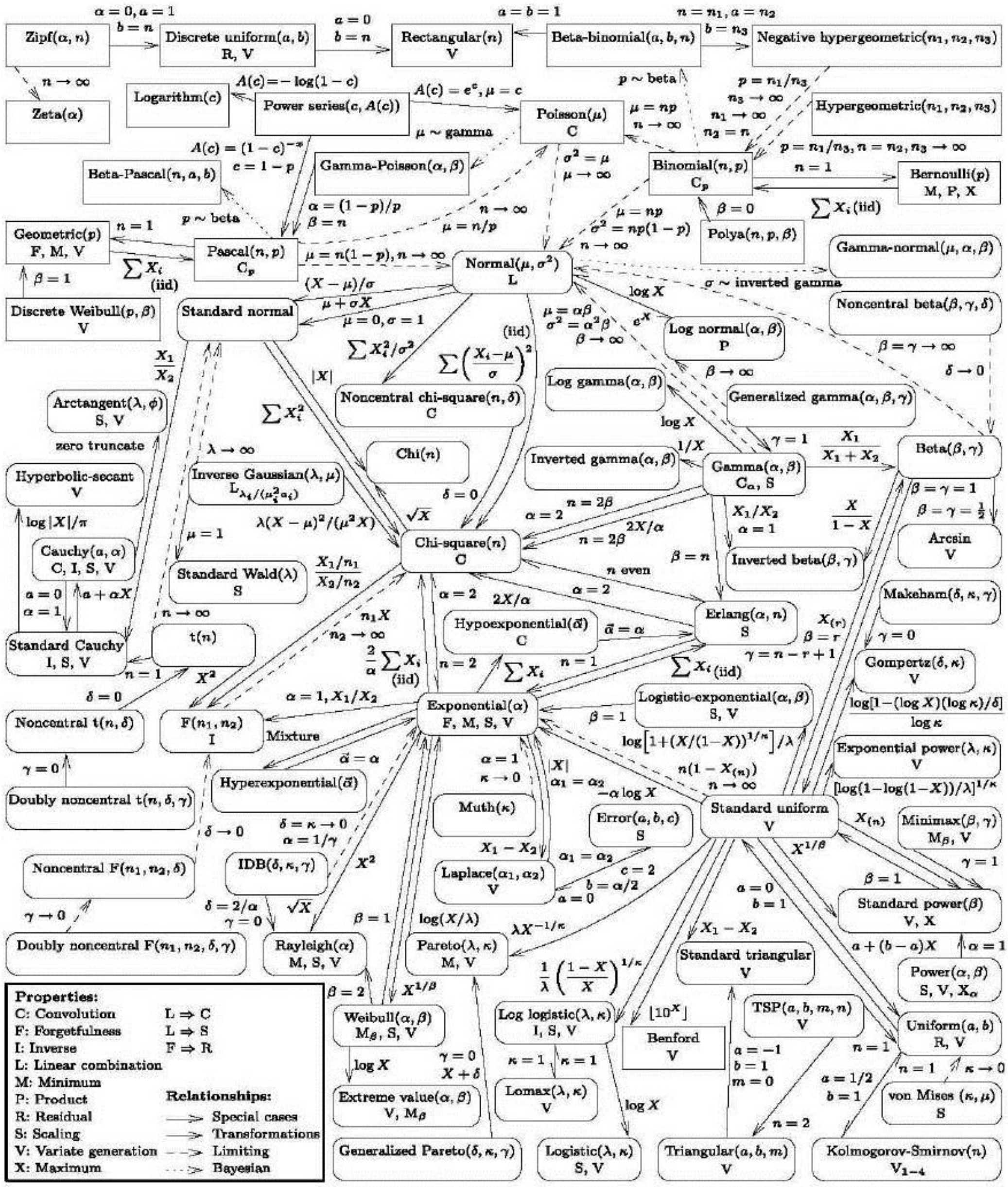
  $L(\mu,\gamma)\sim\gamma L(0,1)+\mu.$
- 图像特征: $F(\mu-x)+F(\mu+x)=1.$
- 回归模型:

$$P_i=\frac{1}{1+\mathrm{e}^{-(a+bx_i)}}\quad\Rightarrow\quad\ln\left(\frac{P_i}{1-P_i}\right)=a+bx_i.$$

- 推广: 多元 Logistic 函数 $y=(1+\mathrm{e}^{-\boldsymbol{\beta x}})^{-1}.$

其它分布

- 超指数分布
- Dirichlet 分布
- 广义 Dirichlet 分布
- 组合 Dirichlet 分布
- 刘维尔分布
- 威布尔分布
- 埃尔朗分布
- 帕累托分布



A.2.3 多维离散型

1 多项分布

$X=(X_1,\cdots,X_n)\sim M(N;p_1,\cdots,p_n).$

$$P(X_1=k_1,X_2=k_2,\cdots,X_n=k_n)=\frac{N!}{k_1!k_2!\cdots k_n!}p_1^{k_1}p_2^{k_2}\cdots p_n^{k_n}.$$

多项分布的边缘分布是二项分布.

$(X_1, X_2, \cdots, X_n) \sim M(N; p_1, p_2, \cdots, p_n) \quad \Rightarrow \quad X_1 + X_2 \sim B(N; p_1 + p_2).$

A.2.4 多维连续型

1 矩形均匀分布

2 二维正态分布

2.1 基础概念

- $X = (X_1, X_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho).$

$$f(x_1, x_2) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)\right].$$

- 当且仅当 $\rho = 0$ 时, X_1 和 X_2 独立.
- 边缘分布 $X_1 \sim N(\mu_1, \sigma_1^2),\ X_2 \sim N(\mu_2, \sigma_2^2).$

2.2 数字特征

- 相关系数 $\mathrm{Corr}(X_1, X_2) = \rho.$
- 协方差 $\mathrm{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2.$
- 期望 $E(X_1X_2) = \mathrm{Cov}(X_1, X_2) + E(X_1)E(X_2) = \rho\sigma_1\sigma_2 + \mu_1\mu_2.$

2.3 其它性质

- 二维正态分布的边缘分布是正态分布.
- 二维正态分布的条件分布是正态分布.

  若 $(X, Y) \sim N(a, b, \sigma_1^2, \sigma_2^2, \rho)$, 则给定 $X = x$ 时 Y 的条件分布为

  $$N(b + \rho\sigma_2\sigma_1^{-1}(x-a),\ \sigma_2^2(1-\rho^2)).$$

- 二维正态分布的边缘分布的和仍为正态分布

  若 $(X_1, X_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则 $Y = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2).$

- 独立的正态分布的联合分布是正态分布.

  正态分布的联合分布不一定是二维正态分布.

- 若 $Y = X_1 + X_2$ 服从正态分布, X_1, X_2 独立, 则 X_1, X_2 也是正态分布. ⭐

3 多元正态分布

3.1 基础概念

- 设 $(X_1, X_2, \cdots, X_n)$ 为 n 元随机变量, 令

  $$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{nn} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & n_{nn} \end{pmatrix},$$

  其中 $\boldsymbol{C}$ 为协方差矩阵, $c_{ij} = \mathrm{Cov}(X_i, X_j) = \rho_{ij}\sigma_i\sigma_j.$

  如果 $(X_1, X_2, \cdots, X_n)$ 的概率密度函数为

  $$f(x_1, x_2, \cdots, x_n) = \frac{\mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}}{(2\pi)^{\frac{n}{2}}|\boldsymbol{C}|^{\frac{1}{2}}}$$

  则称 $(X_1, X_2, \cdots, X_n)$ 是参数为 $\boldsymbol{\mu},\ \boldsymbol{C}$ 的 n 元正态变量.

3.2 数字特征

- 方差: $\mathrm{Var}(X_i) = c_{ii}$.
- 协方差: $\mathrm{Cov}(X_i, X_j) = c_{ij}$.
- 相关系数: $\mathrm{Corr}(X_i, X_j) = \rho_{ij} = \dfrac{c_{ij}}{\sqrt{c_{ii}c_{jj}}}$.
- 数学期望: $E(X_i X_j) = c_{ij} + \mu_1 \mu_2$.

3.3 其它性质

- n 维正态分布的边缘分布是正态分布.
- n 维正态分布的条件分布是正态分布.
- n 维正态分布的边缘分布的和是正态分布.
- n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 服从 n 维正态分布的充要条件是:

$$\forall l_i \in \mathbb{R}\,(i = 1, 2, \cdots, n) : l_1 X_1 + l_2 X_2 + \cdots + l_n X_n \sim N(\mu, \sigma^2).$$

- 若 $Y_1, Y_2, \cdots, Y_m$ 都是 n 维正态分布分量 $X_i\,(i = 1, 2, \cdots, n)$ 的线性函数, 则 $(Y_1, Y_2, \cdots, Y_m)$ 服从 m 维正态分布.
- n 维正态分布各分量相互对立充要条件是它们两两不相关.

4 狄利克雷分布

4.1 基础概念

- $\boldsymbol{X} \sim \mathrm{Dir}(\boldsymbol{\alpha})$.
- Dirichlet 分布又称为多元 Beta 分布, 属于指数族分布.
- 多元 Beta 函数与 Gamma 函数的 **Dirichlet 公式**

$$\begin{aligned} \mathrm{B}(\boldsymbol{\alpha}) = \mathrm{B}(\alpha_1, \alpha_2, \cdots, \alpha_n) &:= \int \cdots \int \prod_{i=1}^{n} x_i^{a_i - 1}\, \mathrm{d}\boldsymbol{x} && \left(\sum_{i=1}^{n} x_i = 1\right) \\ &= \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_n)}{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n)} = \frac{\Gamma(\alpha_1)\cdots\Gamma(\alpha_n)}{\Gamma(a_0)} && \left(a_0 = \sum_{i=1}^{n} a_n\right) \end{aligned}$$

其中 $d\ (d \in \mathbb{N}^+)$ 维积分域是一个开放的 $d-1$ 维正单纯形, 由顶点 $(1, 0, \cdots, 0), (0, 1, \cdots, 0), \cdots, (0, 0, \cdots, 1)$ 围成.

- 概率密度函数

$$\begin{aligned} \mathrm{Dir}(\boldsymbol{X} \mid \boldsymbol{\alpha}) &= \frac{1}{\mathrm{B}(\boldsymbol{\alpha})} \prod_{i=1}^{d} X_i^{\alpha_i - 1} = \frac{\Gamma(\alpha_0)}{\prod_{i=1}^{d} \Gamma(\alpha_i)} \prod_{i=1}^{d} X_i^{\alpha_i - 1} && \left(\alpha_0 = \sum_{i=1}^{d} \alpha_i, d \geq 3\right) \\ &= \frac{\Gamma(\alpha_0)}{\prod_{i=1}^{d} \Gamma(\alpha_i)} \left(\prod_{i=1}^{d-1} X_i^{\alpha_i - 1}\right) (1 - X_1 - \cdots - X_{d-1})^{\alpha_d - 1} && (\|\boldsymbol{X}\| = 1) \end{aligned}$$

其中 $\boldsymbol{\alpha}$ 是无量纲的 **分布参数**, $d \geq 3$ 为随机变量的维度.

备注:

- 上式中的范数指 1-范数而非 ~~2-范数~~.
- Dirichlet 分布的 d 维*支撑集* 同 Dirichlet 公式中的积分域.
- 概率分布记作 $\mathrm{Dir}(\boldsymbol{\alpha})$, 密度函数记作 $\mathrm{Dir}(\boldsymbol{X} \mid \boldsymbol{\alpha})$.
- 向量 $\boldsymbol{X}$ 是 $n-1$ 维, 而 $\boldsymbol{\alpha}$ 是 n 维.

- 对称 Dirichlet 分布
  - 概率密度函数 $\mathrm{Dir}(\boldsymbol{X} \mid \boldsymbol{\alpha}) = \dfrac{\Gamma(d\boldsymbol{\alpha})}{\Gamma(\boldsymbol{\alpha})^d} \displaystyle\prod_{i=1}^{d} X_i^{\alpha_i - 1}$.

- 对称 Dirichlet 分布在每个概率密度相等, 即分布参数 $\boldsymbol{\alpha}$ 在所有维度相同, 取值也被称为 **浓度参数**.
    - 当浓度参数为 1 时, d 维 Dirichlet 分布退化为 $d-1$ 维正单纯形上的均匀分布, 也被称为 **平 Dirichlet 分布**.
    - 当浓度参数大于 1 时, 对称 Dirichlet 分布是一个*集中分布*, 此时浓度参数越大, 概率密度越集中.
    - 当浓度参数小于 1 时, 对称 Dirichlet 分布是一个*稀疏分布*, 此时浓度参数越接近于 0, Giallo密度越稀疏.

- 累积分布函数

$$F(\boldsymbol{b})=\int_{\mathbb{R}_d\cap[0,b)}\mathrm{Dir}(\boldsymbol{X}\mid\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{X},\quad(\boldsymbol{b}\in(0,1])$$

4.2 数字特征

- 众数
    - $M(X_i)=\dfrac{\alpha_i-1}{\alpha_i+\alpha_d-2}(x_d+x_i)$.

      注: $x_d+x_i=1-x_1-\cdots-x_{i-1}-x_{i+1}-\cdots-x_{d-1}$ 与 x_i 无关.
    - $M(X_i)=\dfrac{\alpha_i-1}{\alpha_0-d}$.

      注: 这是所有分量都取到众数时的取值, 是上式的特例.
- 矩 $E\left(\prod\limits_{i=1}^d X_i^{\beta_i}\right)=\dfrac{\Gamma(\alpha_0)}{\Gamma(\alpha_0+\beta_0)}\prod\limits_{i=1}^d\dfrac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)}=\dfrac{\mathrm{B}(\boldsymbol{\alpha}+\boldsymbol{\beta})}{\mathrm{B}(\boldsymbol{\alpha})}$, $\beta_0=\sum\limits_{i=1}^d\beta_i$.
    - 期望 $E(X_i)=\dfrac{\alpha_i}{\alpha_0}$.
    - 方差 $\mathrm{Var}(X_i)=E(X_i^2)-E(X_i)^2=\dfrac{\alpha_i(\alpha_0-\alpha_i)}{\alpha_0^2(\alpha_0+1)}$.
    - 协方差 $\mathrm{Cov}(X_i,X_j)=E(X_iX_j)-E(X_i)E(X_j)=\dfrac{\alpha_i(\alpha_0-\alpha_j)}{\alpha_0^2(\alpha_0+1)}$.

4.3 其它性质

- 相关分布
    - 边缘分布 $p(X_i)=\mathrm{Be}(X_i\mid\alpha_i,\alpha_0-\alpha_i)$.

      备注:
        - 即 Beta 分布, 或 2 维 Dirichlet 分布.
        - '|' 符号类似分号, 与条件概率毫无关系, 上式可写作 $F(x_i;\alpha_i,\alpha_0-\alpha_i)$.
    - 联合分布

$$p(X_i,X_j)=\mathrm{Dir}(X_i,X_j\mid\boldsymbol{\alpha}),\quad\boldsymbol{\alpha}=[\alpha_i,\alpha_j,\alpha_0],\quad i,j\in\{1,2,\cdots,d\}.$$

      即边缘分布 X_i 和 X_j 的联合分布为 3 维 Dirichlet 分布.
- 作为概率分布的性质
    - 共轭性: 多项分布的共轭先验是 Dirichlet 分布 (同等无知原则).
    - 聚合性: 不懂.
    - 中立性

      任意的 $(X_1,X_2,\cdots,X_s)\in\boldsymbol{X}$ 都与归一化后的 $(X_{s+1},\cdots,X_d)\in\boldsymbol{X}$ 相互独立:

$$(X_1,\cdots,X_s)\perp\boldsymbol{X}^*,\quad\boldsymbol{X}^*=\left(\frac{X_{s+1}}{X_{s+1}+\cdots+X_d},\cdots,\frac{X_d}{X_{s+1}+\cdots+X_d}\right),$$
$$p(\boldsymbol{X}^*\mid X_1,X_2,\cdots,X_s)=\mathrm{Dir}(\boldsymbol{\alpha}^*),\quad\boldsymbol{\alpha}^*=(\alpha_{s+1},\alpha_{s+2},\cdots,\alpha_d).$$

    - Dirichlet 是服从 Gamma 分布的 d 维 iid 随机变量 $\boldsymbol{T}=\Gamma(\boldsymbol{T}\mid\boldsymbol{\alpha},1)$ 归一化后的联合分布:

$$
\begin{aligned}
&T_i = \Gamma(T_i \mid \alpha_i, 1), \quad Z_d = \textstyle\sum_{i=1}^d T_i \\
&\boldsymbol{X} = \frac{1}{Z_d}(T_1, T_2, \cdots, T_{d-1}), \\
&p(\boldsymbol{X}) = \mathrm{Dir}(\alpha_1, \alpha_2, \cdots, \alpha_d).
\end{aligned}
$$

- 信息测度

A.3 数列和常数

A.3.1 数列

卡特兰数

Catalan 数又称明安图数.

递归定义

1. $C_0 = C_1 = 1$.
2. $C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k}$.

前几项值: 1, 1, 2, 5, 14, 43, 132, 429, 1439, 4862, 16796...

生成函数

由 $G(x) = \sum_{n=0}^{\infty} C_n x^n$ 知 $G^2(x) = \sum_{n=0}^{\infty} C_{n+1} x^n$, 故

$$
\begin{aligned}
&\begin{cases} G(x) = \sum_{n=0}^{\infty} C_n x^n \\ C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k} \end{cases} &\Rightarrow \quad & G^2(x) = \sum_{n=0}^{\infty} C_{n+1} x^n \\
&\begin{cases} xG^2(x) + 1 = G(x) \\ G(0) = 1 \end{cases} &\Rightarrow \quad & G(x) = \frac{1 - \sqrt{1-4x}}{2x}
\end{aligned}
$$

通项公式

1. $C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{1}{2n+1}\binom{2n+1}{n}$.
2. $C_n = \binom{2n}{n} - \binom{2n}{n-1}$.
3. $C_n = \frac{1}{n+1}\sum_{i=0}^{n}\binom{n}{i}^2$.

证明

1. 由生成函数泰勒展开即得.
2. 由组合数定义即得.
3. 对比 $(1+x)^n\left(1+\frac{1}{x}\right)^n = \frac{(1+x)^{2n}}{x^n}$ 两边系数, 即得 $\sum_{i=0}^{n}\binom{n}{i}^2 = \binom{2n}{n}$.

递推公式: $C_n = \frac{4n-2}{n+1} C_{n-1}$.

证明 由通项公式即得.

例题

1. 满足通项关系 $C_n = \binom{2n}{n} - \binom{2n}{n-1}$ 的场景.
    1. 在 $n \times n$ 网格中, 一开始在 $(0,0)$ 处, 每次可以向上走一格或向右走一格, 在任一时刻, 向右的次数不少于向上的次数, 则合法的路径有 $\binom{2n}{n} - \binom{2n}{n-1} = C_n$ 种.
    2. 有 n 对括号, 则长度为 $2n$ 的括号序列中合法的序列有 C_n 种. (入栈出栈)
    3. 一个圆周上有 $2n$ 个点, 两两配对并连线, 则所有弦不相交的连法有 C_n 种.
2. 满足递归定义 $C_n = \sum\limits_{k=0}^{n-1} C_k C_{n-k-1}$ 的场景.
    1. 把一个 n 层的矩形阶梯分为 n 个矩形的方法有 C_n 种.
    2. 凸 $n+2$ 边形按顶点连线划分为 n 个三角形的方法有 C_n 种.

A.3.2 常数

卡特兰常数

级数定义与积分定义

$$\begin{aligned} G = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2} &= \int_0^1 \frac{\arctan x}{x}\,\mathrm{d}x \\ &= -\int_0^1 \frac{\ln x}{1+x^2}\,\mathrm{d}x = \int_0^{+\infty} \frac{\ln x}{1+x^2}\,\mathrm{d}x \\ &= -\int_0^{\frac{\pi}{4}} \ln\tan x\,\mathrm{d}x = \int_0^{\frac{\pi}{4}} \ln\cot x\,\mathrm{d}x \\ &= 0.915965594177219015054603515\ldots \end{aligned}$$

常用积分

- $\left[0, \frac{\pi}{2}\right]$.
    - 对数与三角
        - 正余弦 (区间再现后相加)
            - $\int_0^{\frac{\pi}{2}} \ln\sin x\,\mathrm{d}x = -\frac{\pi}{2}\ln 2.$
            - $\int_0^{\frac{\pi}{2}} \ln\cos x\,\mathrm{d}x = -\frac{\pi}{2}\ln 2.$
        - 正余切 (卡特兰常数定义)
            - $\int_0^{\frac{\pi}{2}} \ln\tan x\,\mathrm{d}x = -G.$
            - $\int_0^{\frac{\pi}{2}} \ln\cot x\,\mathrm{d}x = G.$
        - $1 \pm$ 正余弦 (由半角公式即得)
            - $\int_0^{\frac{\pi}{2}} \ln(1+\sin x)\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \ln(1+\cos x)\,\mathrm{d}x = 2G - \frac{\pi}{2}\ln 2.$
            - $\int_0^{\frac{\pi}{2}} \ln(1-\sin x)\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \ln(1-\cos x)\,\mathrm{d}x = -2G - \frac{\pi}{2}\ln 2.$
        - $1 +$ 正余切 (分区间利用结论)

            - $\int_0^{\frac{\pi}{2}} \ln(1+\tan x)\,\mathrm{d}x = G + \frac{\pi}{4}\ln 2.$
            - $\int_0^{\frac{\pi}{2}} \ln(1+\cot x)\,\mathrm{d}x = G + \frac{\pi}{4}\ln 2.$
    - 幂与三角 (分布积分用结论)
        - $\int_0^{\frac{\pi}{2}} \frac{x\sin x}{1+\cos x}\,\mathrm{d}x = 2G - \frac{\pi}{2}\ln 2.$
        - $\int_0^{\frac{\pi}{2}} \frac{x\sin x}{1-\cos x}\,\mathrm{d}x = 2G + \frac{\pi}{2}\ln 2.$
        - $\int_0^{\frac{\pi}{2}} \frac{x\cos x}{1+\sin x}\,\mathrm{d}x = -2G + \pi\ln 2.$
        - $\int_0^{\frac{\pi}{2}} \frac{x\cos x}{1-\sin x}\,\mathrm{d}x = +\infty.$
- $\left[0, \frac{\pi}{4}\right].$
    - 对数与三角
        - 正余弦 (相加减后解方程)
            - $\int_0^{\frac{\pi}{4}} \ln\sin x\,\mathrm{d}x = -\frac{1}{2}G - \frac{\pi}{4}\ln 2.$
            - $\int_0^{\frac{\pi}{4}} \ln\cos x\,\mathrm{d}x = \frac{1}{2}G - \frac{\pi}{4}ln2.$
        - $1\pm$ 正余切 (区间再现后展开)
            - $\int_0^{\frac{\pi}{4}} \ln(1+\tan x)\,\mathrm{d}x = \frac{\pi}{8}\ln 2.$
            - $\int_0^{\frac{\pi}{4}} \ln(1-\tan x)\,\mathrm{d}x = \frac{\pi}{8}\ln 2 - G.$
            - $\int_0^{\frac{\pi}{4}} \ln(\cot x+1)\,\mathrm{d}x = G + \frac{\pi}{4}\ln 2.$
            - $\int_0^{\frac{\pi}{4}} \ln(\cot x-1)\,\mathrm{d}x = \frac{\pi}{8}\ln 2.$
        - 正余弦和差 (平方之后二倍角)
            - $\int_0^{\frac{\pi}{4}} \ln(\cos x+\sin x)\,\mathrm{d}x = \frac{1}{2}G - \frac{\pi}{8}\ln 2.$
            - $\int_0^{\frac{\pi}{4}} \ln(\cos x-\sin x)\,\mathrm{d}x = -\frac{1}{2}G - \frac{\pi}{8}\ln 2.$
    - 幂与三角
        - $x\cdot$ 正余切 (分布积分用结论)
            - $\int_0^{\frac{\pi}{4}} x\tan x\,\mathrm{d}x = \frac{1}{2}G - \frac{\pi}{8}\ln 2.$
            - $\int_0^{\frac{\pi}{4}} x\cot x\,\mathrm{d}x = \frac{1}{2}G + \frac{\pi}{8}\ln 2.$
- 其它区间
    - 幂与对数 (三角换元用结论)
        - $\int_0^1 \frac{\ln\left(1+x^2\right)}{1+x^2}\,\mathrm{d}x = -G + \frac{\pi}{2}\ln 2.$
        - $\int_0^1 \frac{\ln\left(1-x^2\right)}{1+x^2}\,\mathrm{d}x = -G + \frac{\pi}{4}\ln 2.$
        - $\int_0^{\frac{\sqrt{2}}{2}} \frac{\ln x}{\sqrt{1-x^2}}\,\mathrm{d}x = -\frac{1}{2}G - \frac{\pi}{4}\ln 2.$
