R1

预备知识(5)

① 定积分的换元法

设函数 $f(x) \in C[a,b]$，单值函数 $x=\varphi(t)$ 满足

1) $\varphi(t) \in C^1[\alpha,\beta]$，$\varphi(\alpha)=a$，$\varphi(\beta)=b$；

2) 在 $[\alpha,\beta]$ 上，$a \le \varphi(t) \le b$

则 $\int_a^b f(x)dx = \int_\alpha^\beta f[\varphi(t)]\varphi'(t)dt$

② $\frac{d}{dx}\int_{\psi(x)}^{\varphi(x)} f(t)dt = f[\varphi(x)]\varphi'(x) - f[\psi(x)]\psi'(x)$

③ $\begin{cases} x=x(u,v) \\ y=y(u,v) \end{cases}$ $dxdy = \frac{\partial(x,y)}{\partial(u,v)}dudv$

④ 泰勒展开：$f(x) = \frac{f(a)}{0!} + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x)$

⑤ 伽玛函数 $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$

1) $\Gamma(1)=1$，$\Gamma(\frac{1}{2})=\sqrt{\pi}$ ； 2) $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$；

3) 当 $n\in N^+$ 时 $\Gamma(n+1)=n!$

$Ga(\alpha,\lambda)$，$E=\frac{\alpha}{\lambda}$，$Var=\frac{\alpha}{\lambda^2}$

伽玛分布

$p(x)=\begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x\ge 0 \\ 0, & x<0 \end{cases}$

$\alpha=1$ 时

$Ga(1,\lambda)=Exp(\lambda)$

$\alpha=\frac{n}{2}$，$\lambda=\frac{1}{2}$ 时

$Ga(\frac{n}{2},\frac{1}{2})=\chi^2(n)$

⑥ 贝塔函数 $B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}dx$

1) $B(a,b)=B(b,a)$

2) $B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

若 $X\sim Ga(\alpha,\lambda)$

则 $aX\sim Ga(\alpha,\frac{\lambda}{a})$

① 三大抽样分布（与其他分布联系，分位数---）（在各随机变量独立前提下）

(1) 卡方分布

设 $X_1, X_2, \cdots X_n$ 独立同分布于 $N(0,1)$，则 $\chi^2 = X_1^2 + \cdots + X_n^2$ 称为自由度为 n 的 χ^2 分布.

$\chi^2(n)$ 分布的密度函数 $p(y)$ 为：

$$p(y) = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}}, \quad y>0$$

$E_{\chi^2(n)} = n$，$Var_{\chi^2(n)} = 2n$

关于 $\chi^2(n)$ 的一条重要定理：

设 $x_1, \cdots x_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本，定义样本均值 $\bar{x}$ 和样本方差 S^2

分别为 $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ 和 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$，则有：

(1) $\bar{x}$ 与 S^2 独立； (2) $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$； (3) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

(2) F分布

$$F \sim F(m,n) = \frac{\left(\frac{\chi^2(m)}{m}\right)}{\left(\frac{\chi^2(n)}{n}\right)} \Rightarrow \left(\frac{1}{F} \sim F(n,m)\right).$$

$F(m,n)$ 的密度函数 $p_F(y)$，

$$p_F(y) = \frac{\Gamma\left(\frac{m+n}{2}\right)\left(\frac{m}{n}\right)^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} y^{\frac{m}{2}-1}\left(1+\frac{m}{n}y\right)^{-\frac{m+n}{2}}, \quad y \geq 0$$

对于 F分布的分位数，满足 $F_\alpha(n,m) = \frac{1}{F_{1-\alpha}(m,n)}$.

关于 F分布的一条重要定理：

设 $x_1, \cdots x_m$ 是来自 $N(\mu_1, \sigma_1^2)$ 的样本，$y_1, \cdots y_n$ 是来自 $N(\mu_2, \sigma_2^2)$ 的样本，且此两种样本相互独立，则 $F = \frac{\left(\frac{S_x^2}{\sigma_1^2}\right)}{\left(\frac{S_y^2}{\sigma_2^2}\right)} \sim F(m-1, n-1)$

(3) t分布

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

t(n)的密度函数为 $p_t(y) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1+\frac{y^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < y < +\infty$

t(n)的相关性质：

(1) n=1时 t分布为标准柯西分布，均值不存在

(2) n>1时 $E_{t(n)} = 0$ ；

(3) n>2时 $Var_{t(n)} = \frac{n}{n-2}$

关于 t分布的一条~~自~~重要定理：

设 $x_1, \cdots, x_n$ 是来自正态分布 $N(\mu,\sigma^2)$ 的一个样本，$\bar{x}$ 与 s^2 分别是该样本的样本均值与样本方差，则有.

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

$$= \frac{\left(\frac{(\bar{x}-\mu)\sqrt{n}}{\sigma}\right)}{\left(\sqrt{\frac{(n-1)s^2}{\sigma^2(n-1)}}\right)}$$

关于 t 分布的分位数， $t_\alpha(n) = -t_{1-\alpha}(n)$

R2

①

概念：试验 ⟶ 样本空间(Ω) ⟶ 事件集类

试验 ⟶ 事件(A) ⟶ 事件集类

⟶ 赋以概率（不同解释）

运算：事件 — 和 — 容斥公式 $\xrightarrow{\text{互斥}}$ 加法公理

事件 — 积 — 乘法法则 $\xrightarrow{\text{独立}}$ 简单公式

概率

全概率公式 ⟶ 贝叶斯公式

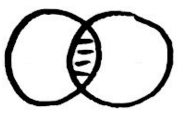
② 事件的记号

补事件：A^c

↓ 对立

$A+B=\Omega$

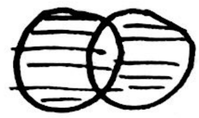
积事件：AB



↓ 互斥

$AB=\phi$

和事件：$A+B$



③ 基本的事件概率运算公式

1)

$$P(A_1+A_2+\cdots+A_n)=\sum_{j=1}^{n}P(A_j)-\sum_{i_1<i_2}P(A_{i_1}A_{i_2})+\cdots+(-1)^{r+1}\sum_{i_1<\cdots<i_r}P(A_{i_1}A_{i_2}\cdots A_{i_r})$$

$$+\cdots+(-1)^{n+1}P(A_1A_2\cdots A_n)$$

e.g. 配对问题

[redacted]

i：至少1人配对成功的概率？

ii：恰有k个人配对成功的概率？

2) 乘法法则

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

$$\overset{独立}{=\!=}P(A_1)P(A_2)\cdots P(A_n)$$

(注①: $P(B)>0$, $P(A|B)\triangleq\dfrac{P(AB)}{P(B)}$ (条件概率的定义)

判断 A、B 是否独立? $\Rightarrow P(A|B)\overset{?}{=}P(A)\Rightarrow P(AB)\overset{?}{=}P(A)P(B)$)

(注②: 容斥公式 $\xrightarrow{互斥}$ 加法公理

乘法法则 $\xrightarrow{独立}$ 简化法则)

3) 全概率公式:

$$P(A)=P(\sum AB_i)=\sum P(AB_i)$$

$$=\sum P(A|B_i)P(B_i)$$

(B_i 相互独立)

贝叶斯公式:

$$P(B_i|A)=\frac{P(AB_i)}{P(A)}$$

$$=\frac{P(A|B_i)P(B_i)}{\sum P(A|B_i)P(B_i)}$$

④ 1) 随机变量：样本空间上的实值函数

- 离散型：随机变量至多可数个取值
- 连续型，......

2) (累积)分布函数：$F(x) \triangleq P(X \leq x)$，$\forall x \in R$ （$P(a<X\leq b)=F(b)-F(a)$）

- 1) 单调非减
- 2) $\lim\limits_{x\to-\infty} F(x)=0$，$\lim\limits_{x\to+\infty} F(x)=1$
- 3) 右连续

3) 概率密度函数（连续型随机变量）

若存在 $f\geq 0$，s.t. $\forall$ 区间 $I \subset R$，有 $P(X\in I)=\int_I f(x)\,dx$.

则称 f 为连续型随机变量 X 的概率密度函数 →
- 1) $\int_{-\infty}^{+\infty} f(x)\,dx=1$
- 2) $P(a<X\leq b)=\int_a^b f(x)\,dx$
- ⋮

（f 不唯一；应用常选择 $f=F'$，F 为累积分布函数）.

4) 期望 {连续型, 离散型} （$E(x)=\int_{-\infty}^{+\infty} x f(x)\,dx$）. 方差 {连续型, 离散型} （$Var(X)=E(X^2)-E^2(X)$ $=\int_{-\infty}^{+\infty}(x-E(X))^2 f(x)\,dx$ ----

⑤ 常用离散分布 I

1) 伯努利分布　$X=\begin{cases}1, & P(X=1)=p\\ 0, & P(X=0)=1-p\end{cases}$，

$Var(X)=p(1-p)$，$E(X)=p$.

2) 二项分布（实例：掷 n 次 1 元硬币）

定义 X = n 次伯努利实验成功次数. $k=0,1,\cdots,n$.

$P(X=k)=C_n^k\,p^k(1-p)^{n-k}$，记为 $X\sim B(n,p)$

3) 泊松分布

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

$Var(X)=\lambda$, $E(X)=\lambda$

(e.g. 时间 $[0,1)$ 内路口事故数 $\begin{cases}\text{时间小区间足够小, st 每一个小区间内至多发生1次} \\ \text{恰发生一次, } p=\frac{\lambda}{n} \\ \text{小区间 相互独立}\end{cases}$)

注: 当 (在二项分布中), p 很小, n 很大, $\lambda=np$ 不太大时

可用 $\lambda=np$ 的泊松分布来近似二项分布

(e.g. 配对问题: 恰有 k 个人拿到自己的帽子 $\begin{cases}\text{指定 } k \text{ 个人拿到} \\ \text{其余 } (n-k) \text{ 个人没拿到.}\end{cases}$)

⑥ 常用连续分布 I ($f(x)$: 概率密度函数 (pdf); $F(x)$: (累积)分布函数 (cdf))

1) 均匀分布

$$f(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b \\ 0, & \text{others}\end{cases},\ 记为\ X\sim U(a,b)$$

(当 $X\sim U(0,1)$ 时 称 X 为随机数) $E(X)=\frac{a+b}{2}$, $Var(X)=\frac{(b-a)^2}{12}$

2) 指数分布

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0 \\ 0, & x\le 0\end{cases}\ (\lambda>0) \quad \left(F(x)=\begin{cases}1-e^{-\lambda x}, & x>0 \\ 0, & x\le 0\end{cases}\right)$$

"无记忆性"

$E(X)=\frac{1}{\lambda}$, $Var(X)=\frac{1}{\lambda^2}$ $\quad P(X>t+s \mid X>s)=P(X>t)$

另: "瞬时失效率"

$$F(x)=1-e^{-\left(\frac{x}{\beta}\right)^{\alpha}},\quad x>0$$

$$\lim_{h\to 0^+}\frac{P(x<X<x+h\mid X>x)}{h}=\cdots=\frac{F'(x)}{1-F(x)}=\lambda \quad \text{or} \quad \frac{\alpha}{\beta^{\alpha}}x^{\alpha-1}$$

3) 正态分布

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad x\in R,\ \text{记作 } X\sim N(\mu,\sigma^2)$$

$E(X)=\mu$， $Var(X)=\sigma^2$

特别地，若 $X\sim N(\mu,\sigma^2)$，令 $Y=\dfrac{X-\mu}{\sigma}$，则 $Y\sim N(0,1)$ ← 标准正态分布

令 $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$，$x\in R$（标准正态分布的概率密度函数）

$\Phi(x)=\int_{-\infty}^{x}\varphi(t)\,dt$（标准正态分布的累积分布函数）

（ $\Phi(x)=1-\Phi(-x)$ ）

（若 $X\sim N(0,1)$，则 $E(X)=E(X^3)=0$.）

$$\left\{E(X^2)=1 \rightarrow \int_{-\infty}^{+\infty}x^2e^{-\frac{x^2}{2}}dx=\sqrt{2\pi}\right.$$

① 随机变量的函数（ $Y=g(X)$ ）

$E(Y)=\int_{-\infty}^{\infty}g(x)f(x)\,dx=\text{———}$ （求期望等数字量）

$F(y)=P(Y\le y)=P(g(x)\le y)=\text{———}$ （求cdf、pdf）

一点重要结论：（利用随机数生成不同分布）

若 X 的连续、严格递增累积分布函数为 $F(x)$，则随机变量 $Y=F(X)\sim U(0,1)$

or

若 $X\sim U(0,1)$，F 为分布函数(连续、严格递增)，则 $Y=F^{-1}(X)\sim F$

③ 常用离散分布 II

1) 多项分布（实例：掷n次六面色子）

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

$(k_1+k_2+\cdots+k_n=N)$（$X_i=k_i$：第i种情况出现k_i次）

④ 常用连续分布 II（$f(x,y)$：联合密度函数）.

1) 均匀分布

$$f(x,y) = \begin{cases} \frac{1}{(b-a)(d-c)}, & a<x<b,\ c<y<d \\ 0, & \text{others} \end{cases}$$

另：圆盘上的均匀分布？

2) 二元正态分布

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

注1：二元正态分布的边际分布

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

注2：二元正态分布的条件分布

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

$$= \frac{p(x,y)}{p_Y(y)}$$

注3: $-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)$

$$= -\frac{1}{2(1-\rho^2)}[x, y]\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$= -\frac{1}{2}[x, y]\ A^T A\begin{bmatrix} x \\ y \end{bmatrix}$$

其中 $A = \frac{1}{\sqrt{1-\rho^2}}\begin{bmatrix} 1 & -\rho \\ 0 & \pm\sqrt{1-\rho^2} \end{bmatrix}$ 或 $A = \frac{1}{\sqrt{1-\rho^2}}\begin{bmatrix} -1 & \rho \\ 0 & \pm\sqrt{1-\rho^2} \end{bmatrix}$

(10) 条件分布 及条件期望 (此处 $p(x,y)$ 通 $f(x,y)$)

1) (把 $P(X\le x \mid Y=y)$ 看成是 $\lim\limits_{h\to 0} P(X\le x \mid y\le Y\le y+h)$).

定义: 对一切使 $P_Y(y)>0$ 的 y, 给定 $Y=y$ 条件下 X 的条件分布函数为

$$F(x|y) = \int_{-\infty}^{x}\frac{p(u,y)}{p_Y(y)}du$$

X 的条件密度函数为

$$p(x|y) = \frac{p(x,y)}{p_Y(y)}$$

注1: 由条件密度函数得联合概率密度函数

$$p(x,y) = p(x|y)\,p_Y(y) = p(y|x)\,p_X(x)$$

进而 $$p(x|y) = \frac{p(y|x)\,p_X(x)}{\int_{-\infty}^{\infty} p(y|x)\,p_X(x)\,dx}$$ 贝叶斯公式

$$p_Y(y) = \int_{-\infty}^{\infty} p(y|x)\,p_X(x)\,dx$$ 全概率公式

(回到条件密度函数)

2) （在 $X=x$ 的条件下求 Y 的期望）

定义: $E(Y|X=x)=\begin{cases}\sum_j y_j P(Y=y_j|X=x), & \text{离散.}\\ \int_{-\infty}^{\infty} x p(x|y)\,dx, & \text{连续}\end{cases}$

注1: (重期望公式)（$E(Y|X=x)$ 是 x 的函数）.

$$E(X)=E\big(E(X|Y)\big)$$

$$=\begin{cases}\sum_j E(X|Y=y_j)P(Y=y_j) & \text{. 离散}\\ \int_{-\infty}^{\infty} E(X|Y=y)\,p_Y(y)\,dy & \text{, 连续.}\end{cases}$$

（要求在一个取值于很大范围上的指标 X 的 $E(X)$，
可以先找一个与 X 相关的量 Y，用 Y 将大范围划分为若干小范围，
求小范围上的 $E(X|Y=y)$，再利用重期望公式即可）（与贝叶斯公式：
$P(y|x)$ 与 $E(Y|X=x)$ ）

注1.1:

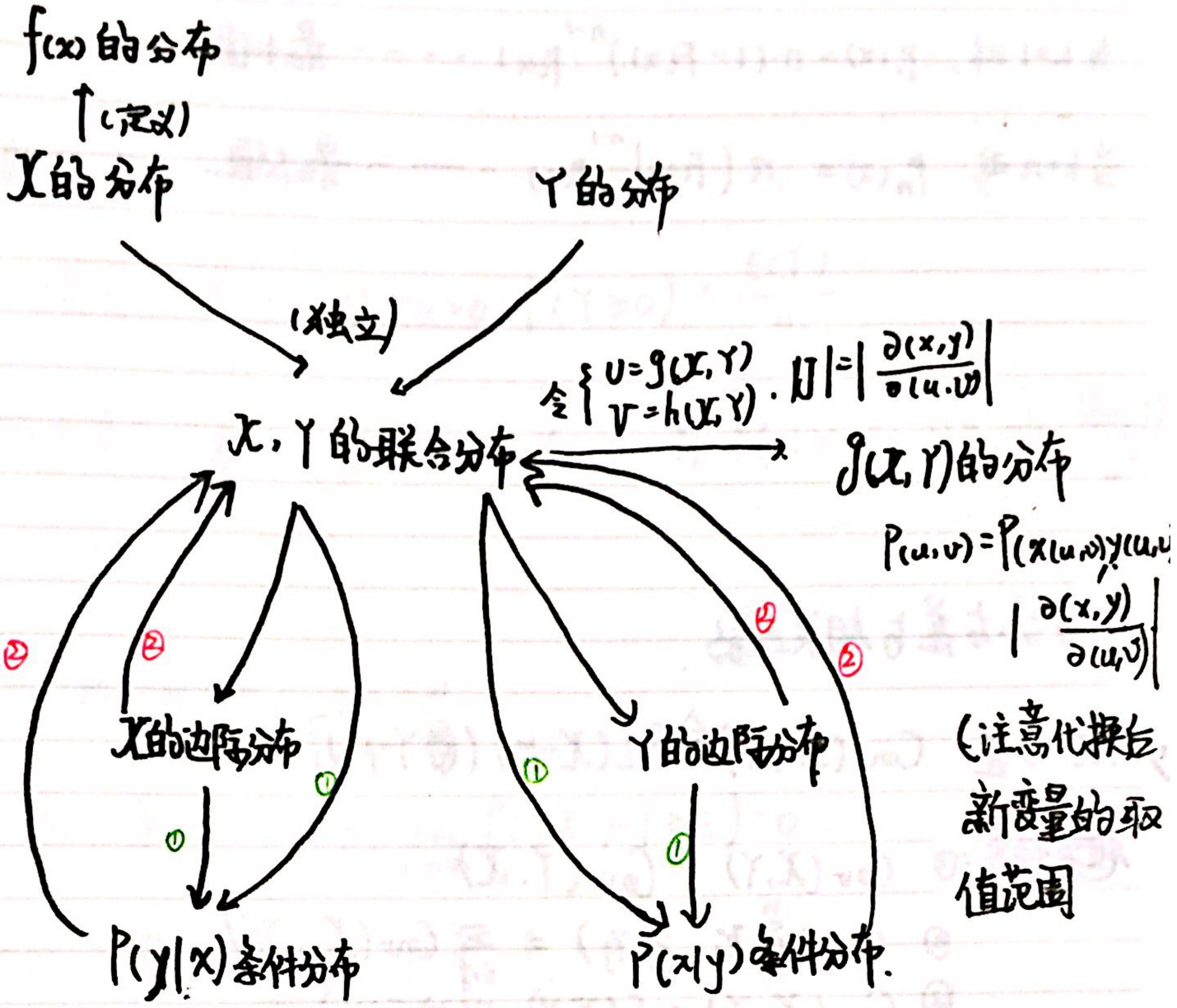
设 $X_1, X_2, \cdots$ 为一列独立同分布的随机变量. 随机变量 N 只取整数值，且 N 与 $\{X_n\}$ 独立，则

$$E\left(\sum_{i=1}^{N} X_i\right)=E(X_1)E(N)$$

注2:

$$E\left[(Y-g(X))^2\right]\geq E\left[(Y-E(Y|X))^2\right]$$

⑪ 联合分布、边际分布、条件分布的联系 (X、Y 为随机变量)。



①: $P(y|x) = \dfrac{P(x,y)}{P_X(x)}$ ②: $P(x,y) = P(y|x)\cdot P_X(x)$

(基于各密度分布函数，积分得累积分布函数)

⑫ 常用分布性质

1) n 个独立同分布的指数变量和为伽玛分布:

$$\underbrace{Exp(\lambda) * Exp(\lambda) \cdots * Exp(\lambda)}_{n个} = Ga(n,\lambda)$$

2) 设总体 X 的密度函数为 $p(x)$, 分布函数为 $F(x)$, $x_1, x_2, \cdots x_n$ 为样本, 则第 k 个次序统计量 $x_{(k)}$ 的密度函数为

$$p_k(x) = \frac{n!}{(k-1)!(n-k)!}(F(x))^{k-1}(1-F(x))^{n-k}p(x)$$

特别地，

当 $k=1$ 时，$p_1(x) = n\cdot(1-F(x))^{n-1}p(x)$ ---- 最小值

当 $k=n$ 时 $p_n(x) = n(F(x))^{n-1}p(x)$ ---- 最大值.

⑬ 协方差与相关系数.

1) 协方差: $Cov(X,Y) \triangleq E[(X-\mu_1)(Y-\mu_2)]$

性质: ① $Cov(X,Y) = Cov(Y,X)$

② $Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i,j} Cov(X_i, Y_j)$

③ $Cov(X,Y) = E(XY) - E(X)E(Y)$

④ $Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + \sum_{i\neq j} Cov(X_i, X_j)$

2) 相关系数.

$$\rho = Corr(X,Y) \triangleq \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = E\left(\frac{X-\mu_1}{\sigma_1}\cdot\frac{Y-\mu_2}{\sigma_2}\right)$$

注①: 设 U,V 为随机变量，则 $E^2(UV) \le E(U^2)E(V^2)$

(利用 $E[(V-tU)^2] \ge 0,\ \forall t\in R$)

注②: $\rho=\pm 1 \Leftrightarrow \exists b\ st\ P(Y=\pm\frac{\sigma_2}{\sigma_1}X+b)=1$

注③: 据说，在正态分布下，不相关 $\Leftrightarrow$ 独立

注(3.1)： $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. $\Rightarrow \rho=\mathrm{Corr}(X,Y)$.

⑭ 1) 大数定律的 Pre：

(1) 马尔可夫不等式：$Y\geq 0$，则 $\forall a>0$，$P(Y\geq a)\leq \dfrac{E(Y)}{a}$.

(2) 契比雪夫不等式：若 $Var(Y)$ 存在，则 $\forall a>0$，有 $P(|Y-E(Y)|\geq a)\leq \dfrac{Var(Y)}{a^2}$.

2) 大数定律：

· 若 $X_1,\cdots X_n$ 独立同分布，$E(X_i)=\mu$，

则
- (弱大数定律)：$\forall \varepsilon>0$，$\lim\limits_{n\to\infty} P(|\bar{X}-\mu|\geq\varepsilon)=0$.
- (强大数定律)：$P(\lim\limits_{n\to\infty}\bar{X}=\mu)=1$.

3) 中心极限定理：

若 $X_1,\cdots X_n$ 独立同分布，$E(X_i)=\mu$，$Var(X_i)=\sigma^2$，

则 $\lim\limits_{n\to\infty} P\left(\dfrac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\,\sigma}\leq x\right)=\Phi(x)$

注：当 $n\to\infty$ 时 $\bar{X}\overset{近似}{\sim}N(\mu,\dfrac{\sigma^2}{n})$

⑭ 参数估计
- 点估计
  - 矩估计 $\hat{\lambda}$
  - 极大似然估计 λ^*
- 区间估计
  - 小样本
  - 大样本.

14.1 矩估计 ☆ (基于大数定律)(常被采用)

设总体具有已知的概率函数 $p(x;\theta_1,\cdots,\theta_k)$, $(\theta_1,\cdots,\theta_k)\in\Theta$ 为未知参数.

假定总体的k阶原点矩存在 ($\forall j$, $0<j<k$, μ_k 存在)

若 $\theta_1,\cdots\theta_k$ 能表示成 $\mu_1,\cdots\mu_k$ 的函数 $\theta_j=\theta_j(\mu_1,\cdots\mu_k)$,

则 θ_j 的矩估计为

$$\hat{\theta}_j=\theta_j(a_1,\cdots a_k),\ j=1,\cdots k, \text{其中 } a_j=\frac{1}{n}\sum_{i=1}^{n}x_i^j \text{ 为样本k阶原点矩.}$$

进一步. 若 $\eta=g(\theta_1,\cdots\theta_k)$.

则 $\hat{\eta}=g(\hat{\theta}_1,\cdots\hat{\theta}_k)$

(将待估量以总体矩表示,代入样本矩,得矩估计)

注: 尽量采用低阶矩给出未知参数的估计.

14.2 极大似然估计 ☆

Only 似然函数: 一组观测值(样本) $(x_1,\cdots x_n)$ 出现的概率

$$L(\theta)=L(x_1,\cdots x_n;\theta)$$

求极大似然估计, 即求 $\theta^*=\theta^*(x_1,\cdots x_n)$ 使 $L(\theta)$ 达到最大.

注: 求偏导得零点后检验是否为最大值点!

14.3 区间估计（注：说明取下α分位数）

给定 $\alpha \in (0,1)$，对于任意所有可能的取值θ，都有

$$P(\hat{\theta}_1 \le \theta \le \hat{\theta}_2) \ge 1-\alpha \text{，} \hat{\theta}_i = \hat{\theta}_i(X_1, \cdots X_n), i=1,2.$$

则称 $[\hat{\theta}_1, \hat{\theta}_2]$ 为θ的置信水平为 $100(1-\alpha)\%$ 的置信区间.

关键：寻找枢轴量 $G = G(x_1, \cdots x_n, \theta)$，其分布不依赖于待估计参数，

通过 "$c \le G \le d \Leftrightarrow \hat{\theta}_1 \le \theta \le \hat{\theta}_2$" 求出 $\hat{\theta}_1$ 和 $\hat{\theta}_2$

$P(c \le G \le d) = 1-\alpha$ 可算，即 c, d 可确定.

① $X_i \sim N(\mu, \sigma^2)$，$\sigma^2$ 已知，对μ作区间估计：（大样本）

$\because \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ $\boxed{\therefore \frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \sim N(0,1)}$.

记N(0,1)的下α分位数为 z_α.

$$\therefore P\left(z_{\frac{\alpha}{2}} \le \frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \le z_{1-\frac{\alpha}{2}}\right) = P\left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}} \le \mu \le \bar{x} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$\Rightarrow$ μ的区间估计为 $\left[\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{x} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right]$

② $X_i \sim N(\mu, \sigma^2)$，$\sigma^2$ 未知，对μ作区间估计（小样本）

$\because \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ $\quad \therefore \frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \sim N(0,1)$

且 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

$\therefore \dfrac{\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma}}{\sqrt{\frac{\frac{(n-1)S^2}{\sigma^2}}{n-1}}} = \boxed{\frac{\sqrt{n}(\bar{x}-\mu)}{S} \sim t(n-1)}$，记 $t(n)$ 的下α分位数为 $t_\alpha(n)$

$$\therefore P\left(t_{\frac{\alpha}{2}}(n-1) \le \frac{\sqrt{n}(\bar{x}-\mu)}{S} \le t_{1-\frac{\alpha}{2}}(n-1)\right) = P\left(\bar{x} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \le \mu \le \bar{x} - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right) = 1-\alpha$$

$\Rightarrow \mu$ 的区间估计为 $\left[\bar{X}-\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1), \bar{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right]$

③ $X_i \sim N(\mu,\sigma^2)$，μ、σ^2 未知，对 σ^2 作区间估计：（小样本）

$\because \boxed{\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)}$

$\therefore$ 记 $\chi^2(n)$ 的下 α 分位数为 $\chi^2_{\alpha}(n)$

$\therefore P\left(\chi^2_{\frac{\alpha}{2}}(n-1) \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{1-\frac{\alpha}{2}}(n-1)\right)$

$= P\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right) = 1-\alpha$

$\Rightarrow \sigma^2$ 的区间估计为 $\left[\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right]$

④ $X_i \sim N(\mu_1,\sigma^2)$，$Y_i \sim N(\mu_2,\sigma^2)$，$\mu_1,\mu_2,\sigma^2$ 未知，X、Y 独立，对 $\mu_1-\mu_2$ 作区间估计
$(X_1,\cdots X_n)$ $(Y_1,\cdots Y_m)$ （小样本）

$\because \bar{X}-\bar{Y}-(\mu_1-\mu_2) \sim N\left(0, \frac{\sigma^2}{n}+\frac{\sigma^2}{m}\right)$.

$\therefore \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim N(0,1)$

又 $\because \frac{(n-1)S_A^2}{\sigma^2}+\frac{(m-1)S_B^2}{\sigma^2} \sim \chi^2(m+n-2)$

$\therefore \frac{\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\frac{1}{\sigma}\sqrt{\frac{(n-1)S_A^2+(m-1)S_B^2}{m+n-2}}} = \boxed{\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}\cdot\sqrt{\frac{(n-1)S_A^2+(m-1)S_B^2}{m+n-2}}} \sim t(m+n-2)}$

$\therefore P\left(t_{\frac{\alpha}{2}}(m+n-2) \le \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}\cdot\sqrt{\frac{(n-1)S_A^2+(m-1)S_B^2}{m+n-2}}} \le t_{1-\frac{\alpha}{2}}(m+n-2)\right) = 1-\alpha$

$\Rightarrow \mu_1-\mu_2$ 的区间估计为 — — 、 —

⑤ $X \sim Exp(\lambda)$，对λ作区间估计（小样本）

$\because X_i \sim Exp(\lambda) = Ga(1, \lambda)$

$\Rightarrow X_1 + \cdots + X_n \sim Ga(n, \lambda)$.

$\therefore \bar{x} \sim Ga(n, n\lambda) \Rightarrow \boxed{2n\lambda\bar{x} \sim Ga(n, \tfrac{1}{2}) = \chi^2(2n)}$.

$\therefore P\left(\chi^2_{\frac{\alpha}{2}}(2n) \leq 2n\lambda\bar{x} \leq \chi^2_{1-\frac{\alpha}{2}}(2n)\right) = P\left(\frac{\chi^2_{\frac{\alpha}{2}}(2n)}{2n\bar{x}} \leq \lambda \leq \frac{\chi^2_{1-\frac{\alpha}{2}}(2n)}{2n\bar{x}}\right) = 1-\alpha$

⇒ λ的区间估计为 $\left[\frac{\chi^2_{\frac{\alpha}{2}}(2n)}{2n\bar{x}}, \frac{\chi^2_{1-\frac{\alpha}{2}}(2n)}{2n\bar{x}}\right]$

⑥ 对于二项分布 $B(n,p)$，对 p 作区间估计

（$\frac{\bar{x}-p}{\sqrt{\frac{p(1-p)}{n}}} \overset{近似}{\sim} N(0,1)$，用样本的$\bar{x}$估计$p(1-p)$）

对于 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim (\mu_2, \sigma_2^2)$，X, Y独立，$\sigma_1^2, \sigma_2^2$未知，对 $\mu_1 - \mu_2$ 作区间估计

（$\frac{(\bar{x}-\bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}} \overset{近似}{\sim} N(0,1)$，用$S_1^2, S_2^2$估计$\sigma_1^2, \sigma_2^2$）

对于 X，μ, σ^2 未知，对于 μ 作区间估计

（$\frac{(\bar{x}-\mu)}{\frac{\sigma}{\sqrt{n}}} \overset{近似}{\sim} N(0,1)$）

$\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \overset{近似}{\sim} N(0,1)$

（大样本）

（$\bar{x}$ 总体均值，一个随机变量；$\bar{x}$ 样本均值，被观测到的）

14.4

⑮ 假设检验.

15.1 临界值检验

1) 提出 H_0: - - - - - vs H_1: - - - - - (H₁往往是感兴趣的) ~~[illegible]~~ {可靠的，基于经验认识的

2) 确定检验统计量及分布 ⟶ 拒绝域形式（由 H_1 决定）⟶ 给定 α

3) 由"$\boxed{在 H_0 成立下}$，拒绝 H_0 的概率 $\leq \alpha$" ⇒ '拒绝域'

4) 收集样本

5) 得出结论 .

Notes: $\bar{X}, \bar{x}, \mu, \mu_0$ 的区别；→ $\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \sim N(0,1)$ 而非 $\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma} \sim N(0,1)$

P(I) 是在 H_0 成立的前提下计算的；

H_1 是感兴趣的假设

Tips: 各分布下 α 分位数的关系：

$N(0,1)$: $z_\alpha = -z_{1-\alpha}$

$\chi^2(n)$: ?

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

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

15.2 p值检验

1) 提出 H_0: - ______ vs H_1: - ______

2) 确定检验统计量及分布

原假设为真时观测值（$\bar{x}$）及比观测值更极端的观测出现的概率

确定"极端"的形式 $\Rightarrow$ p值计算方式

给定α

3) 收集样本，计算检验统计量的值及p值

4) 得出结论 ($\begin{cases} \alpha \geq p, & \text{拒绝 } H_0 \\ \alpha < p, & \text{不拒绝 } H_0 \end{cases}$).

Note: 当p值很小时即可拒绝H_0，当p值很大时即可接受H_0；

对于不同的样本，有不同的 p值；

e.g. 检验统计量：$\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma}$，观测值：$\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}$