概统期末复习.

概率

Lect 01 概率基本概念及事件运算 6.8

古典概型

样本和样本空间.

事件的运算. Venn图. 示性函数　　互不相容 A B　　对立 $\boxed{A\ \bar{A}}$

$\boxed{\text{线上}}$ 两个著名例子 (to be filled) 一.4

布丰投针: $p=\frac{2l}{\pi d}$　与线相交的概率(一组平行线,针长为l) $p=\frac{\int_0^{\pi}\frac{l}{2}\sin\theta d\theta}{\frac{d}{2}\pi}$

贝特朗奇论: 圆内随机取弦, $\geq\sqrt{3}$的概率? $\frac{1}{4}$(看中心). $\frac{1}{3}$(固定一点). $\frac{1}{2}$(与圆心距离)

Lect 02 条件概率和独立性

Def (条件概率)

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

△ 三个重要公式:

1. 乘法公式 $P(AB)=P(B)P(A|B)$

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

2. 全概率公式

设$B_1\cdots B_n$为样本空间Ω的一个分割,即$B_1\cdots B_n$互不相容且$\bigcup B_i=\Omega$.

若$P(B_i)>0$. 则对任一事件A, 有 $P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$

3. 贝叶斯公式. 应用: B前发生. A后发生

设$B_1\cdots B_n$为Ω的一个分割 若$P(A)>0$. $P(B_i)>0$. 则 $P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n}P(B_j)P(A|B_j)}$

e.g. 确诊率问题. A: 化验呈阳性 B. 患病.

Def (独立性) $P(AB)=P(A)P(B)$. 则A.B独立. 否则不独立/相关.

不要求真的独立. 只要计算满足就可以!

Def. (互斥/互不相容) $P(AB)=0$.

Def. (相关系数). 后面有、 $r(A,B)=\frac{P(AB)-P(A)P(B)}{\sqrt{P(A)(1-P(A))P(B)(1-P(B))}}$　　独立: 0. 相等: 1. $\in[0,1]$.

这是随机事件, 是随机变量的特例. (可理解仅有发生:1. 不发生:0)

$\boxed{\text{线上}}$ 二.3 (10'30"–15'20") 三个事件独立　　二.4 应用

三个事件独立: 两两独立且 $P(ABC)=P(A)P(B)P(C)$

可满足① 不满足②　e.g. 四面体 涂红黄蓝 三色.

A. B. C 表示 将其投一次, 包含X色.

应用: 是否性别歧视? 女生只申请难录取专业 ⇒ 加权入学率.

Monty Hall 换门? 换. $\frac{2}{3}$

Lect03 随机变量(一)

Def (随机变量) 定义在样本空间 Ω 上的函数. $X: \Omega \to \mathbb{R}$.

e.g. $x = X(\omega)$, $\omega \in \Omega$. $X(\omega)$ 一般简记为 X.

Def (分布函数) $F(x) = P(X \le x)$ 为 X 的分布函数，或 X 服从 $F(x)$. 记为 $X \sim F(x)$.

$\lim\limits_{x\to-\infty} F(x) = 0$. $\lim\limits_{x\to+\infty} F(x) = 1$ $\quad P(a < X \le b) = F(b) - F(a)$ $\quad F_X(x)$.

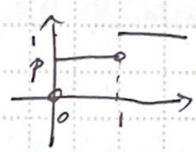
Def (离散型随机变量的分布列).

X	$x_1 \cdots x_n \cdots$
P	$p_1 \cdots p_n \cdots$

$X \sim \begin{pmatrix} x_1 & \cdots & x_n & \cdots \\ p_1 & \cdots & p_n & \cdots \end{pmatrix}$

X 可能取值是有限多或可列多个

两点分布. Bernoulli分布. $\begin{pmatrix} a_0 & a_1 \\ p & q \end{pmatrix}$ $\Rightarrow p+q=1$ $\quad F(x) = \begin{cases} 0 & x < a_0 \\ p & a_0 \le x < a_1 \\ 1 & x \ge a_1 \end{cases}$



① 二项分布: $X \sim b(n,p)$

伯努利试验重复 n 次. $p_k = \binom{n}{k} p^k (1-p)^{n-k}$. $k = 0,1,\cdots,n$.

② 泊松分布: $X \sim P(\lambda)$. 描述稀有事件.

对给定常数 $\lambda > 0$. $p_k = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$, $k = 0,1,2,\cdots$

Taylor e^{λ} 在0处展开. $\sum\limits_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda} \Rightarrow \sum p_k = 1$

Thm 二项分布 $b(n,p)$ 当 p 很小 n 很大时, $b(n,p)$ 与 $P(np)$ 非常接近. 泊松定理

若 n 重伯努利试验中, 发生概率为 p_n (与 n 有关!). 若 $n \to \infty$ 时 $np_n \to \lambda$.

则 $\lim\limits_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$. $\quad$ Proof. P20.

③ 几何分布: $X \sim Ge(p)$

伯努利试验第一次成功时的次数. $P(X=k) = (1-p)^{k-1} \cdot p \quad k=1,2,\cdots$

Thm. 几何分布的无记忆性 $\quad P(X > s+t \mid X > s) = P(X > t)$

逆命题: 若取正整数值的随机变量 X, 满足 $P(X>s+t \mid X>s) = P(X>t)$, 则 $X \sim Ge(p)$.

证: P29.

④ 负二项分布. $X \sim Nb(r,p)$

第 r 次成功时的试验次数. $P(X = r+k) = \binom{r+k-1}{k} (1-p)^k p^r \quad k = 0,1,2,\cdots$

或 $P(X=k) = \binom{k-1}{r-1} (1-p)^{k-r} p^r \quad k = r, r+1, \cdots$

练习 四.1 负二项分布.

$$P(X=r+k) = C_{r+k-1}^{r-1} \underbrace{(1-p)^k}_{= q^k} p^r = \underbrace{C_{-r}^{k}}_{负!} p^r (-q)^k.$$

$(q = 1-p)$.

Lect04 随机变量(二)　　6.9

Thm(分布函数性质).

(1) 单调性. $\forall x_1 < x_2$ $F(x_1) \le F(x_2)$

(2) 有界性 $\forall x \in \mathbb{R}$. 有 $0 \le F(x) \le 1$. 且 $F(-\infty) = \lim\limits_{x\to-\infty} F(x) = 0$. $F(+\infty) = \lim\limits_{x\to+\infty} F(x) = 1$

(3) 右连续性 $F_X(x)$是关于x的右连续函数

$\forall x_0 \in \mathbb{R}$. 有 $\lim\limits_{x\to x_0^+} F(x) = F(x_0)$. 即 $F(x_0+0) = F(x_0)$

Def(连续型随机变量)

对$F(x)$. 若存在非负可积的$p(x)$. $x\in\mathbb{R}$. 使$\forall x \in \mathbb{R}$ 有 $F(x) = \int_{-\infty}^{x} p(t)\,dt$. 则$X$为连~.

$p(x)$为X的**概率密度函数** 且 $\int_{-\infty}^{+\infty} p(x)\,dx = 1$.　$F'(x) = p(x)$.

① 均匀分布. $-\infty < a < b < +\infty$. $X \sim U(a,b)$?

$$p(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & \text{其它} \end{cases} \qquad F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x < b \\ 1 & x \ge b \end{cases}$$

柯西分布: $p(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$

② 指数分布: $X \sim Exp(\lambda)$, $\lambda > 0$.

$$p(x) = \begin{cases} 0 & x \le 0 \\ \lambda e^{-\lambda x} & x > 0 \end{cases} \qquad F(x) = \begin{cases} 0 & x \le 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases}$$

Thm.(无记忆性) $\forall s > 0$. $t > 0$. $P(Y > s+t \mid Y > s) = P(Y > t)$

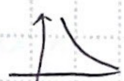
Thm.(与几何分布的联系). $X \sim Exp(\lambda)$, $\lambda > 0$, $Y \sim Ge(1-e^{-\lambda})$

$P(Y = k) = P(k-1 \le X < k)$. $k = 1, 2, \cdots$

$$(1-e^{-\lambda k}) - (1-e^{-\lambda(k-1)}) = (e^{-\lambda})^{k-1}(1-e^{-\lambda}) = (1-p)^{k-1} p.$$

③ 幂律. 二八法则/帕雷托法则.

$p(x) = c x^{-\alpha}$　$x \ge x_{min}$



eg. 社交网络.

④ 正态分布: $X \sim N(\mu, \sigma^2)$ → 分散程度

标准正态分布 $N(0,1)$

$$p(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \xrightarrow{\frac{X-\mu}{\sigma} \sim N(0,1)} \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$F(x)$　　$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\,dt$ 可查表知.

△已知X的分布, 求$Y = g(X)$的分布.

1. 离散.

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

$\xRightarrow{Y = g(X)}$

Y	$g(x_1)$	$\cdots$	$g(x_n)$	$\cdots$
P	p_1	$\cdots$	p_n	$\cdots$

2. 连续 $Y = g(X) \Rightarrow F_Y(y) = P(Y \le y) = P(g(X) \le y)$　　P28例子

△既非连续也非离散的随变.

↓有$p(x)$　↓可列.

$\frac{1}{2}$　$\frac{1}{2^2}$　$\frac{3}{2^2}$

$\boxed{\text{在线}}$ 三.3. 分布函数性质与特例(6'45-14'12)　剩下三等分, 取中间.

康托尔分布(不离散不连续). $C_1 = (\frac{1}{3}, \frac{2}{3})$ $C_2 = (\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9})$　$\frac{1, 3, 5, 7}{2^3}$

Lect 05 随机变量的数字特征

Def (期望) $E(X)=\begin{cases}\sum_{i=1}^{\infty} x_i p(X=x_i) & 离散\\ \int_{-\infty}^{+\infty} x p(x)dx & 连续\end{cases}$

P19 柯西分布.
不绝收. E不存在.
Var也不存在(定义有E)

E存在条件: 绝对收敛. (条级换序后变, 不应该有多种E)

- $E(cX)=cE(X)$.
- $E(X_1+X_2+\cdots+X_n)=E(X_1)+\cdots+E(X_n)$

Def (方差)

$Var(X)=E((X-E(X))^2)$, 或$D(X)$.

- $Var(aX+b)=a^2Var(X)$.

常用: $Var(X)=E(X^2)-E(X)^2$

平方的期望 - 期望的平方.

Def (标准差)

$\sigma(X)=\sqrt{Var(X)}$. $\Rightarrow$ 注意! $\sigma(aX+b)=a\sigma(X)$ 不是a^2!

1. 二项分布. $X\sim b(n,p)$ $E(X)=np$ $Var(X)=np(1-p)$ 用独立推?
2. 泊松分布. $X\sim P(\lambda)$. $E(X)=\lambda$ $Var(X)=\lambda$.
3. 几何分布 $X\sim Ge(p)$. $E(X)=\frac{1}{p}$ $Var(X)=\frac{1-p}{p^2}$ ☆.
4. 均匀分布 $X\sim U(a,b)$ $E(X)=\frac{a+b}{2}$ $Var(X)=\frac{1}{12}(b-a)^2$ ☆
5. 指数分布 $X\sim Exp(\lambda)$. $E(X)=\frac{1}{\lambda}$ $Var(X)=\frac{1}{\lambda^2}$

☆ 指数分布算积分.

$\int_0^{+\infty}5x^2e^{-3x}dx=\frac{5}{3}\int_0^{+\infty}x^2\,3e^{-3x}dx=\frac{5}{3}E(X^2)$, 其中$X\sim Exp(3)$

$E(X^2)=Var(X)+(E(X))^2=\frac{1}{3^2}+(\frac{1}{3})^2=\frac{2}{9}\Rightarrow$ 积分 $=\frac{2}{9}\cdot\frac{5}{3}=\frac{10}{27}$

6. 正态分布 $X\sim N(\mu,\sigma^2)$ $E(X)=\mu$. $Var(X)=\sigma^2$

Thm. 期望与方差的性质

1. $E((X-c)^2)$ 在 $c=E(X)$ 时达到最小
2. 切比雪夫不等式 $\forall\varepsilon>0$. $P(|X-E(X)|\geq\varepsilon)\leq\frac{Var(X)}{\varepsilon^2}$ P21

Def (原点矩. 中心矩)

$E(X^k)$ k阶原点矩. $E((X-E(X))^k)$ k阶中心矩

e.g. 验血问题 P24 e.g. 算E(X). P29 ✓

$\boxed{在线}$. 六. 快排. (39'46")

冒泡: $\frac{n(n-1)}{2}$ 快排: 最好 $n\log n$ 最坏 $\frac{n(n-1)}{2}$

比较次数期望: $2n\ln n+O(n)$

设排序后为 $y_1\cdots y_n$. 定义 X_{ij} $(1\leq i<j\leq n)$. 若 y_i, y_j 比较过则为1, 否则为0. $E(X)=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}E(X_{ij})$

~~$X_{ij}=1$~~ y_i-y_j 比较过 $\Leftrightarrow$ y_i或y_j在$\{y_i, y_{i+1}\cdots y_j\}$中第一个被选为基准元素

$\frac{2}{j-i+1}$.

Lect06 多维随机变量

6.10

分布函数 F
密度函数 p
分布列（也是 p）

Def.（n 维随机变量）

$X(w)=(X_1(w),\cdots X_n(w))$ 其中 $X_i(w)$ 为定义在同一样本空间 $\Omega=\{w\}$ 上的…

Def.（n 维随变的联合分布函数）

$F(x_1,x_2,\cdots,x_n)=P\{X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n\}$

P3. 会面问题

Thm.（联合分布函数性质）

单调性（分别对 x,y）、有界性、右连续性 $F(x+0,y)=F(x,y)$、非负性 $\to P(a<X<b,\ c<Y<d)=F(b,d)-F(a,d)-F(b,c)+F(a,c)\ge 0$

Def.（二维离散随机变量的联合分布列）

$X\backslash Y$	y_1	$\cdots$	$y_n\cdots$
x_1			
$\vdots$			
x_n			
$\vdots$			

$\sum_i\sum_j p_{ij}=1$

Def.（边缘分布）（离散·二维）

$\begin{pmatrix} x_1 & \cdots & x_n & \cdots \\ p(X=x_1) & \cdots & p(X=x_n) & \cdots \end{pmatrix}$ 称为 X 的边缘分布列. $p(X=x_i)=\sum_j p_{ij}$

Def.（二维连续型随机变量）

有二元非负 $p(x,y)$. s.t. $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)\,du\,dv$，则 (X,Y) 为~.

$p(x,y)$ 联合密度函数. $p(x,y)=\frac{\partial^2}{\partial x\partial y}F(x,y)$

Def.（边际/边缘分布函数. 密度函数） 积分技巧 $Exp(\lambda)$

$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}p_X(u)\,du$. $p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\,dy$

Def.（随机变量的独立性）

若对 $\forall x_1\cdots x_n$. 有 $F(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}F_{X_i}(x_i)$. 则称 $X_1,\cdots X_n$ 相互独立.

离散型等价定义：$p(X_1=x_1\cdots X_n=x_n)=\prod_{i=1}^{n}P(X_i=x_i)$

连续型 - - - - : $p(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}p_{X_i}(x_i)$

Thm.（独立随机变量的性质）

☆ X,Y 互相独立 则 $E(XY)=E(X)E(Y)$ 乘 $\xrightarrow{\text{可推}}$ $Var(X+Y)=Var(X)+Var(Y)$ 加

E 加不需独立性

△常见多元分布.

多项分布. $(X_1,X_2,\cdots,X_m)\sim M(n,p_1,\cdots,p_m)$

E 试验有 m 个事件概率 p_i, $\sum p_i=1$. 重复 n 次后 X_i 表示事件 i 发生了 n_i 次.

$P(X_1=k_1,\cdots,X_m=k_m)=\frac{n!}{k_1!\cdots k_m!}p_1^{k_1}\cdots p_m^{k_m}$

n 维均匀分布 $(X_1,\cdots,X_n)\sim U(D)$. $D\subseteq R^n$. 有界. 体积为 $S_D=\int\!\!\int dx_1\cdots dx_n$.

$p(x_1\cdots x_n)=\begin{cases}\frac{1}{S_D} & (x_1,\cdots,x_n)\in D\\ 0 & \text{其它}\end{cases}$

二维正态分布. $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$ $\mu_1,\mu_2\in R$. $\sigma_1.\sigma_2>0$. $|\rho|\le 1$.

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

Thm. 边际密度为一元正态分布. $X\sim N(\mu_1,\sigma_1^2)$ $Y\sim N(\mu_2,\sigma_2^2)$

但逆命题不成立!! 边际正态, 联合未必二元正态.

多元正态 P33-34

二元正态分布性质补充.

ρ 为 X.Y 的相关系数. $E(XY)=\mu_1\mu_2+\rho\sigma_1\sigma_2$. $r(X,Y)=\frac{E(XY)-E(X)E(Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}=\rho$.

$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ 则 $(aX+bY,cX+dY)$ 也服从二元正态 (但 ρ 不一定为0)
线性组合.

对比一元正态可加性. $X\sim N(\mu_1,\sigma_1)$ $Y\sim N(\mu_2,\sigma_2)$, $X+Y\sim N(\mu_1+\mu_2,\sigma_1+\sigma_2)$

6.10

Lect 07 协方差.相关系数.条件分布.

Def (协方差)

$Cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E(XY)-E(X)E(Y)$ 两种形式!

Thm (协方差性质)

$Cov(X,a)=0$. $Cov(X,Y)=\overset{\text{交换律}}{Cov(Y,X)}$. ☆ $Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)$ 分配律

线性组合: $Cov(c_1X+a,\ c_2Y+b)=c_1c_2Cov(X,Y)$.

$Var(X\pm Y)=Var(X)+Var(Y)\pm 2Cov(X,Y)$. (X.Y独立时, $Cov(X,Y)=0$. 直接相加)

$[Cov(X,Y)]^2\leq Var(X)\cdot Var(Y)=\sigma_X^2\cdot\sigma_Y^2$. 柯西不等式 $(\int f(x)g(x)dx)^2\leq\int f^2(x)dx\int g^2(x)dx$

Def (相关系数) ⇒ 反映线性关系

$Corr(X,Y)=Cov(\frac{X}{\sigma_X},\frac{Y}{\sigma_Y})=\frac{Cov(X,Y)}{\sigma_X\cdot\sigma_Y}=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}\in[-1,1]$.

>0 正相关 <0 负相关 =0 不相关

不相关并不是独立! 例: $X\sim N(0,1)$, $Y=X^2$, $Cov(X,Y)=E(X^3)-E(X)E(X^2)=0$.

Thm. $Corr(X,Y)=\pm1$ 的充分必要条件是 X、Y 间几乎处处有线性关系. 即存在常数 a.b. s.t. $P(Y=aX+b)=1$. P12

Thm. $\min_{a,b}E[Y-(aX+b)]^2=E[Y-(\hat{a}X+\hat{b})]^2=Var(Y)(1-corr(X,Y)^2)$. $\hat{a}=\frac{Cov(X,Y)}{Var(X)}$

Thm. (二元正态分布性质)

二元正态分布的边缘分布是一元正态分布, 但逆命题不成立.

☆二元正态分布 不相关 ⟺ 独立

Thm. (泊松分布. 二项分布的可加性)

① 泊松 $X_1\sim P(\lambda_1)$ $X_2\sim P(\lambda_2)$ ⋯ $X_m\sim P(\lambda_m)$ 且互相独立, 则 $X_1+\cdots+X_m\sim P(\lambda_1+\cdots+\lambda_m)$

② 二项: $X_1\sim b(n_1,p)$ $X_2\sim b(n_2,p)$ ⋯ $X_m\sim b(n_m,p)$ 且互相独立, 则 $X_1+\cdots+X_m\sim B(n_1+\cdots+n_m,p)$

Def (条件分布)

① 离散: $P(X=x_i|Y=y_j)=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{p_{\cdot j}}$

② 连续 $p(x|y)=\frac{p(x,y)}{p_Y(y)}$

Def (全概率公式)

$p_Y(y)=\int_{-\infty}^{+\infty}p_X(x)\,p(y|x)\,dx$.

Def (贝叶斯)

$p(x|y)=\frac{p_X(x)p(y|x)}{\int_{-\infty}^{+\infty}p_X(x)p(y|x)dx}$

Von. Neumann 取舍原则 P30-31.

$\boxed{\text{连线}}$ 7.1. 独立随机变量和的分布 (1'49"-12'41")

$Z=X+Y$. $f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)dx$.

$X_1\cdots X_n\sim Exp(\lambda)$. $S_n=\sum X_i$. 则 $p_{S_n}(z)=\begin{cases}\lambda^n z^{n-1}e^{-\lambda z}/(n-1)! & z\geq0\\ 0 & z<0.\end{cases}$

Lect 08 条件期望与随机变量函数分布

Def (条件分布期望)

$$X|Y=y \sim \begin{cases} 取值 x_i: & P(X=x_i|Y=y) \quad 离散 \\ 取值 x: & p(x|y) \quad 连续 \end{cases}$$

$$E(X|Y=y) = \begin{cases} \sum_i x_i P(X=x_i|Y=y) & 离散 \\ \int_{-\infty}^{\infty} x p(x|y)\, dx & 连续 \end{cases}$$ $\Rightarrow E(X|Y)$ 是一个随机变量.

Def (重期望公式) 例: P7-8

$E(X) = E(E(X|Y))$ 以 y 的各种可能性分类, 算 x 的期望.

Thm (瓦尔德方程)

$X_1 \cdots X_n$ 独立同分布随变, N 为一个取整数值的随变且与 $\{X_k\}$ 独立.

那么 $E(\sum_{i=1}^{N} X_k) = E(X_1)E(N)$

Proof $E(\sum_{i=1}^{N} X_k) = E(E(\sum_{i=1}^{N} X_k | N)) = \sum_{n=1}^{\infty} P(N=n) E(\sum_{i=1}^{N} X_k | N=n)$

$= \sum_{n=1}^{\infty} P(N=n) \cdot n \cdot E(X_1) = E(X_1) \cdot E(N)$.

例: 买煎饼. 排队人数 $Y \sim P(\lambda)$. 每个人买数量 $X \sim Ge(p)$. 所有人共买多少的期望.

Thm (非线性的最小二乘最佳预报)

二元随变 (X, Y). D(X). D(Y) 存在, 令 $\phi(X) = E(Y|X)$. 则 $E(Y-\phi(X))^2 = \min_{\psi} E(Y-\psi(X))^2$.

P12-13.

Thm.

设 n 维随变 $X = (X_1, X_2, \cdots, X_n)$ 的密度函数为 $p_X(x_1, x_2, \cdots, x_n)$.

n 元函数 $\{g_i(x_1, x_2, \cdots, x_n)\}_{i=1}^{n}$ ⇐ 每维都是关于 $x_1 \cdots x_n$ 的函数. 满足条件.

(1) 存在唯一的反函数 $x_i(y_1, \cdots y_n)$ 即存在方程组 $g_i(x_1, \cdots, x_n) = y_i$ 的唯一实数解

(2) $g_i(x_1 \cdots x_n)$ 与 $x_i(y_1, \cdots, y_n)$ 都连续

(3) 存在连续偏导数 $\frac{\partial x_i}{\partial y_j}$, $\frac{\partial g_i}{\partial x_j}$. 记 $J = \det\left(\begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{bmatrix}\right)$

则 n 维随变 $Y = (Y_1, \cdots, Y_n)^T = (g_1(X_1 \cdots X_n), \cdots, g_n(X_1, \cdots, X_n))^T$

的密度函数为 $p_Y(y_1, \cdots, y_n) = \begin{cases} p_X(x_1, \cdots, x_n) \cdot |J| & 当 y=g(x) 有解时 \\ 0 & 其它 \end{cases}$

① 若不唯一反函数则须相加 e.g. P23.

② 若仅有 $m < n$ 个 g_i, 可补充 $g_j(x_1 \cdots x_n) = x_j$ $(m+1 \le j \le n)$ $p_Y(y_1, \cdots, y_m) = \iint_{-\infty}^{+\infty} p_{Y'}(y_1, \cdots y_n) dy_{m+1} \cdots dy_n$

e.g. P26.

Thm. (二元正态随变性质)

① $X_1 \cdots X_n$ 相互独立, $X_k \sim N(\mu_k, \sigma_k^2)$, $k=1,2,\cdots,n$ 则 $X_1 + \cdots + X_k \sim N(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2)$

② $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ 则 $aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2)$

且 $(a_1X + b_1Y, a_2X + b_2Y)$ 也服从二元正态分布.

Def (最大值, 最小值的分布)

$Y = \max(X_1 \cdots X_n)$. $p_Y(y) = \underset{\text{n个数都可能是max}}{n} \cdot p_X(y) \cdot [F_X(y)]^{n-1} \Leftarrow F_Y(y) = [F_X(y)]^n$

在线 八.4. 全期望公式(下) 16'45''

$E(T)=E(X_1)E(N)$　　$E(T^2)=Var(X_1)E(N)+E(X_1)^2\cdot E(N^2)$

$Var(T)=E(T^2)-E(T)^2=Var(X_1)E(N)+E(X_1)^2 Var(N)$

Thm (大数定律)

伯努利大数定律：S_A是n次伯努利试验中A发生的次数. $\lim\limits_{n\to\infty}P\{|\frac{S_A}{n}-p|<\varepsilon\}=1$.

Proof. $P(|\frac{S_A}{n}-p|<\varepsilon)\leq 1-\frac{Var(\frac{S_A}{n})}{\varepsilon^2}=1-\frac{\frac{p(1-p)}{n}}{\varepsilon^2}\xrightarrow{n\to\infty}1$.

大数定律：随变序列$\{X_n\}$. 若有$\lim\limits_{n\to\infty}P\{|\frac{1}{n}\sum X_k-\frac{1}{n}\sum E(X_k)|<\varepsilon\}=1$. 则$\{X_n\}$服从~.

依概率收敛：$\{X_n\}$随变序列, X为一个随变. 若对$\forall\varepsilon>0$. $\lim\limits_{n\to\infty}P(|X_n-X|<\varepsilon)=1$. 则称

$\{X_n\}$依概率收敛于X. 记作 $X_n\xrightarrow{P}X$.

切比雪夫大数定律 设$X_1\cdots X_n\cdots$是两两不相关的随变序列, 方差有界.

则对$\forall\varepsilon>0$. 有$\lim\limits_{n\to\infty}P\{|\frac{1}{n}\sum X_k-\frac{1}{n}\sum E(X_k)|<\varepsilon\}=1$.

马尔科夫. 辛钦大数定律 P4 证明? 　$\frac{1}{n^2}Var(\sum\limits_{i=1}^{n}X_k)\to 0$. / 独立同分布, $E(X_i)=\mu$.

$P(|\frac{X_1+\cdots+X_n}{n}-\mu|<\varepsilon)\to 1$

Thm (中心极限定理)

$\{X_n\}$独立同分布. $E(X_i)=\mu$. $Var(X_i)=\sigma^2$. 则对固定的y有$\lim\limits_{n\to\infty}P(\frac{X_1+\cdots+X_n-n\mu}{\sigma\sqrt{n}}\leq y)=\Phi(y)$

$X_1+\cdots+X_n\sim N(\sum E(X_k),\sum Var(X_k))$.

$X_n\sim B(n,p)$时: $\lim\limits_{n\to\infty}P(\frac{X_n-np}{\sqrt{npq}}\leq x)=\Phi(x)$. $\forall x\in R$.

对比统计中抽样定理 (针对正态总体)

$\overline{X}\sim N(\mu,\frac{\sigma^2}{n})$

△ 整数的校正.

统计　　若有二阶矩则可证(见下)

Var. 但没有正态证明.

Lect 01 统计学基本概念

Def. 总体. 样本. P22

Def 统计量 P25 $X_1\cdots X_n$为取自某总体的样本, $T=T(X_1,\cdots,X_n)$(不含有任何未知参数的样本函数). 即称为统计量.

Def 样本均值 $\overline{x}=\frac{1}{n}\sum x_i$. $\Rightarrow$ $\sum\limits_{i=1}^{n}(x_i-\overline{x})=0$, $c=\overline{x}$使$\sum\limits_{i=1}^{n}(x_i-c)^2$最小

☆ 若总体分布为$N(\mu,\sigma^2)$, 则 $\overline{x}\sim N(\mu,\frac{\sigma^2}{n})$

若总体分布非正态或未知, $E(X)=\mu$, $Var(X)=\sigma^2$. 则 $\overline{x}\overset{\text{渐近分布}}{\sim}N(\mu,\frac{\sigma^2}{n})$ $\Rightarrow$ 中心极限定理

Def 样本方差 $S^2=\frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\overline{x})^2$ 注意 $\frac{1}{n-1}$, 减的是样本均值, 不是μ!

Thm 若总体有二阶矩. 即$E(X)=\mu$. $Var(X)=\sigma^2$, $X_1\cdots X_n$为样本

则 $E(\overline{X})=\mu$. $Var(\overline{X})=\frac{\sigma^2}{n}$. $E(S^2)=\sigma^2$ 无偏!

Def 样本k阶原点矩 $a_k=\frac{1}{n}\sum\limits_{i=1}^{n}x_i^k$

中心 $b_k=\frac{1}{n}\sum\limits_{i=1}^{n}(x_i-\overline{x})^k$

简单随机样本：随机性、相互独立。 次序统计量P31

△三大统计分布.

卡方 $X_1, X_2, \cdots X_n$ 独立同分布服从 $N(0,1)$. 则

$Y = X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi_n^2$ 或 $\chi^2(n)$. 称自由度为 n 的 χ^2 分布.

具可加性. $\chi^2(n) + \chi^2(m) \sim \chi^2(n+m)$

$E(Y) = n$

$Var(Y) = 2n$.

t $X_1 \sim N(0,1)$ $X_2 \sim \chi_n^2$. 互相独立.

$t = \frac{X_1}{\sqrt{X_2/n}} \sim t(n)$ 称为自由度为 n 的t分布.

关于 $x=0$ 对称

n 大时与标准正态分布接近.

F $X_1 \sim \chi_m^2$, $X_2 \sim \chi_n^2$ 相互独立

$F = \frac{\frac{X_1}{m}}{\frac{X_2}{n}} = \frac{nX_1}{mX_2} \sim F(m,n)$ 称自由度为 m 与 n 的 F 分布.

Thm. 统计抽样定理

$X_1, X_2 \cdots X_n$ 为来自 $N(\mu, \sigma^2)$ 总体的样本. 均值和方差为 $\overline{X}$ 和 S^2. 则

(1) $\overline{X}$ 与 S^2 相互独立 (2) $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$ (3) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

$\boxed{\text{在线}}$ 十三.4. 三种重要的统计分布和分位数.

t分布: n小时衰减缓慢 ⇒ 重尾

Lect 02 参数点估计的方法与评价

利用样本 $X_1 \cdots X_n$ 估计整体分布的参数 θ. $\hat{\theta} = \theta(X_1 \cdots X_n)$.

$\hat{\theta}$ 构造不一定唯一, 但需满足一定合理性.

Def. (矩估计): 以样本矩近似理论矩

用样本矩替换理论矩(原点矩、中心矩等), 解方程得近似. (尽量低阶便于计算)

Def. (极大似然估计)

以最大概率解释样本数据. (相对于其它参数)

设总体概率函数为 $p(x;\theta)$, $X_1 \cdots X_n$ 样本的联合概率函数是关于 θ 的函数. 用

$L(\theta; x_1, x_2, \cdots, x_n)$ 表示, 简记为 $L(\theta) = \prod_k p(x_k; \theta)$, 称为样本的似然函数.

若有 $L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta)$, 则 $\hat{\theta}$ 是 θ 的极大似然估计. 对数似然函数 $\ln(L(\theta))$

① Def (相合性) 评价标准: 估计量随样本增加而逼近参数真实值

若 $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$, $\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$. 则 $\hat{\theta}_n$ 是 θ 的相合估计. (n 为样本量)

② Def (无偏性) 保证没有系统偏差 $E(\hat{\theta}) = \theta$. 则称 $\hat{\theta}$ 是 θ 的无偏估计

③ Def (有效性) 希望估计围绕参数波动的幅度越小越好.

若 $\hat{\theta}_1$. $\hat{\theta}_2$ 均为无偏估计 若对 $\forall \theta \in \Theta$ 有 $Var(\hat{\theta}_1) \le Var(\hat{\theta}_2)$. 则 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

$X_1 \cdots X_n$ 为 $N(\mu,\sigma^2)$ 样本，矩估计 $\hat{\mu}=\overline{X}$，$\hat{\sigma}^2=S^2$.

最大似然估计 $\hat{\mu}=\overline{X}$. $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$

④Def (均方误差) $MSE(\hat{\theta})=E(\hat{\theta}-\theta)^2=Var(\hat{\theta})+(E(\hat{\theta})-\theta)^2$

若 $\hat{\theta}$ 无偏，则 $MSE(\hat{\theta})=Var(\hat{\theta})$ 与有效性同.

Lect03 参数区间估计

用一个区间估计未知参数

Def. 设 $X_1 \cdots X_n$ 是来自总体 $X\sim F(x;\theta)$ 的样本，$\theta\in\Theta\subset R^1$ 为未知参数

$I(X_1,\cdots,X_n)$ 为随机区间，由样本值完全确定. 称该区间为 θ 的一个置信水平为

$1-\alpha$ 的区间估计，指 $P_\theta(\theta\in I(X_1,\cdots,X_n))\geqslant 1-\alpha$. $\forall\theta\in\Theta$.

总体 $N(\mu,\sigma^2)$，σ^2 已知，μ 未知. 求 $\boxed{\mu}$ 的 $1-\alpha$ 置信区间.

$\overline{X}\sim N(\mu,\frac{\sigma^2}{n}) \Rightarrow \frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}\sim N(0,1). \Rightarrow -u_{1-\frac{\alpha}{2}}\leqslant\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}\leqslant u_{1-\frac{\alpha}{2}}$

$\Rightarrow \mu\in[\overline{X}-\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}, \overline{X}+\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}]$.

u: $N(0,1)$ 的分位数 $P(X\leqslant u_\alpha)=\alpha$

还有 $\chi^2_\alpha(n)$. $t_\alpha(n)$. $F_\alpha(m,n)$

总体 $N(\mu,\sigma^2)$，σ^2 未知，μ 未知. 求 $\boxed{\mu}$ 的 $1-\alpha\cdots$

$\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}\sim N(0,1)$. $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$ $\xrightarrow[\text{独立}]{\text{相互}}$ $\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}\cdot\sqrt{\frac{\sigma^2/(n-1)}{(n-1)S^2}}=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$.

$\Rightarrow \mu\in[\overline{X}-\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1), \overline{X}+\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1)]$

总体 $N(\mu,\sigma^2)$. σ^2 未知，μ 未知，求 $\boxed{\sigma^2}$ 的 $1-\alpha\cdots$

$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1) \Rightarrow \sigma^2\in[\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}]$

Behrens-Fisher: $X_1\cdots X_m$ 来自 $N(\mu_1,\sigma_1^2)$ $Y_1\cdots Y_n$ 来自 $N(\mu_2,\sigma_2^2)$. μ_1 μ_2 σ_1^2 σ_2^2 未知.

① 求 $\mu_2-\mu_1$ 的 $1-\alpha\cdots$

无解. 当 $\sigma_1^2=\sigma_2^2=\sigma^2$ 时. $\overline{X}-\overline{Y}\sim N(\mu_1-\mu_2,(\frac{1}{m}+\frac{1}{n})\sigma^2)$. $\frac{(m-1)S_X^2+(n-1)S_Y^2}{\sigma^2}\sim\chi^2(m+n-2)$ ...

② 求 σ_1^2/σ_2^2 的 $1-\alpha\cdots$

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

③ 当①问题 n 够大时，可用正态分布近似解.

$\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\sigma_X^2/m+\sigma_Y^2/n}}\sim N(0,1) \Rightarrow \frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{S_X^2/m+S_Y^2/n}}\sim N(0,1)$

总体 $b(1,p)$. $X_1,\cdots,X_n$. 求 p 的 $\cdots$

n较大时，有近似分布 $\overline{X}\sim N(p,\frac{p(1-p)}{n})$.

近似求解得 $p\in[\overline{X}-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}, \overline{X}+u_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}]$

$\boxed{\text{左例}}$ 十二. 4. 伪随机数和随机模拟.

Lect04 假设检验

① 建立假设 $H_0: \theta\in\Theta_0$ vs $H_1: \theta\in\Theta_1$ 不一定互斥

② 选择检验统计量，给出拒绝域形式 W

③ 选择显著性水平 $\alpha := \max\{P(\text{拒绝}H_0 \mid H_0\text{为真})\} = \max\{P_\theta(X\in W), \theta\in\Theta_0\}$.

④ 给出 W 具体范围.

正态总体均值的假设检验

(1) $H_0: \mu\le\mu_0$ vs $H_1: \mu>\mu_0$ 单侧检验（偏大异常）

(2) $H_0: \mu\ge\mu_0$ vs $H_1: \mu<\mu_0$ 单……（-小-）

(3) $H_0: \mu=\mu_0$ vs $H_1: \mu\ne\mu_0$ 双……

u检验（方差已知） $u=\frac{\bar{x}-\mu_0}{\sqrt{\sigma^2/n}}=\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}\sim N(0,1)$

(1) $W: \bar{x}>\mu_0+\frac{\sigma}{\sqrt{n}}u_{1-\alpha}$. (2) $\bar{x}<\cdots$ (3) $\bar{x}>\mu_0+\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}$ 或 $\bar{x}<\mu_0-\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}$

t检验（方差未知） $t=\frac{\bar{x}-\mu}{S/\sqrt{n}}=\frac{\sqrt{n}(\bar{x}-\mu_0)}{S}\sim t(n-1)$

(1) $W: \bar{x}>\mu_0+\frac{S}{\sqrt{n}}t_{1-\alpha}(n-1)$ (2) $\bar{x}>\cdots$ (3) $\bar{x}>\mu_0+\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1)$ 或……

Def.（P值） 从每次具体观测 ⇒ p值.（这次 α 是多少能接受?）

原假设成立条件下，检验统计量出现在比观测值更异常范围的概率的最大值.

e.g. $H_0: \mu\ge10$ vs $H_1: \mu<10$. $\bar{x}=9.3$

$p=P(\bar{x}\le9.3)=\cdots=\Phi(-)=\cdots$ $\mu\ge10$ 里 max 取 $\mu=10$.

$H_0: \mu=35$ vs $H_1: \mu\ne35$. $\bar{x}=36.5$ 总体 $\sim N(\mu,3^2)$, $n=36$.

$p=P(|\bar{X}-35|\ge1.5 \mid \mu=35)=1-(\Phi(3)-\Phi(-3))=0.0027$.

Lect05 拟合优度检验，似然比检验

Def.（势函数/功效函数）

$$g(\theta)=P_\theta(X\in W)=\begin{cases}\alpha(\theta) & \theta\in\Theta_0 \quad \text{第一类错误 拒真}\\ 1-\beta(\theta) & \theta\in\Theta_1 \quad \text{第二类不错. 不接假}\end{cases}$$

若对 $\forall\theta\in\Theta_0$ 有 $g(\theta)\le\alpha$，则称该检验是（显著性）水平为 α 的显著性检验

若 $\sup\limits_{\theta\in\Theta_0}\beta(\theta)=\alpha$（$\le$），则称一个功效函数为 $\beta(\theta)$ 的检验是真实水平为 α 的检验（水平为 α……）

Def.（检验函数）

$\phi(x)=\begin{cases}1, & x\in W\\ 0, & x\notin W.\end{cases}$ 给出 $\phi(x)$ ⟺ 给出拒绝域 W ⟺ 给出一个检验.

检验的势函数 $\beta(\theta)=P_\theta(X\in W)=E_\theta(\phi(X))$

$\phi(x)$ 仅取 0,1：非随机化检验；否则为随机化检验

e.g. 女士品茶 让离散型成为真实水平为 α 的检验

$\phi(x)=\begin{cases}1, & x=4\\ r, & x=3\\ 0, & x<3\end{cases}$ 解得 $r=0.156$.

Def (拟合优度检验) / χ^2检验.

总体服从离散分布. 进行n次独立观测. K个取值出现频次为N_i ($\sum N_i = k$).

则 $X = \sum_{i=1}^{k} \frac{(N_i - np_i)^2}{np_i}$ 近似服从 $\chi^2(k-1)$. 自由度K-1. (可理解为$\sum N_i = K$限制-1)

注意分母 np_i

若还需要通过样本估计S个参数, 则X分布近似服从 $\chi^2(k-s-1)$.

列联表检验 (独立性检验). 行合计 列合计.

$$Y = \sum_{i=1}^{s}\sum_{j=1}^{t} \frac{\left(n_{ij} - n\cdot\frac{c_i}{n}\cdot\frac{d_j}{n}\right)^2}{n\cdot\frac{c_i}{n}\cdot\frac{d_j}{n}} = \sum_{i=1}^{s}\sum_{j=1}^{t}\frac{(n\cdot n_{ij} - c_i d_j)^2}{n c_i d_j} \sim \chi^2((s-1)(t-1)).$$