一、事件及其概率

事件的运算 $(\bigcup A_k)^c = \bigcap A_k^c$ $(\bigcap A_k)^c = \bigcup A_k^c$ $\bigcup A_k = \sum A_k \bigcap_1^{k-1} A_i$

古典概型：有限＋等可能

1、$n \to r$

(1) 可重复排列 n^r，不可重复排列 A_n^r

(2) 不可重复组合 C_n^r

(3) 可重复组合 C_{n+r-1}^r

2、$r \to n$

(1) 可容纳、不限 n^r

(2) 可容纳、限 1 A_n^r

(3) 不可容纳，不限 C_{n+r-1}^r

(4) 不可容纳，限 1 C_n^r

3. 多组组合 $n \to r_1, \dots, r_k$

$(a_1 + \dots + a_k)^n \longrightarrow \frac{n!}{r_1! \cdots r_n!}$

4、$n_1, \dots, n_k$ 排列 $\frac{n!}{n_1! \cdots n_k!}$

几何概型：无限＋等可能

$P(\cdot)$：加法定理 $P(\bigcup A_k) = \sum P(A_k) + \sum_2^n (-1)^{k-1} \sum_{1 \le i_1 < \dots < i_k \le n} P(A_{i_1} \cdots A_{i_k})$

次可加性 $P(\bigcup A_k) \le \sum P(A_k)$

上、下连续性

$P(\cdot | A)$： $P(B|A) = P(AB) / P(A)$

$P(\bigcap A_k) = \prod P(A_k | \bigcap_0^{k-1} A_i)$

$B_1, \dots, B_n$ 完备：$P(A) = \sum P(A|B_i) P(B_i)$

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

$A_1, \dots, A_n$ 相互独立：$P(A_{i_1} \cdots A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$

$P(\widehat{A}_1 \cdots \widehat{A}_n) = P(\widehat{A}_1) \cdots P(\widehat{A}_n)$，$\widehat{A}_i = A_i + \overline{A}_i$

$P(B|\bar{A}) = \frac{P(A\bar{B})}{P(\bar{A})}$ $\quad \cap A_n^C = (\cup A_n)^C$

二、随机变量及其分布

cdf $F(x)=P(X\le x)$ $\nearrow$ pmf $\quad P(X=x_k)=p_k$

cdf $\searrow$ pdf $\quad F(x)=\int_{-\infty}^{x} f(t)\,dt$

随机变量函数的分布

$P(X=x_k)=p_k$, $P(Y=g(x_k))=\sum\limits_{i:g(x_i)=g(x_k)} p_i$

$X\sim f(x)$, $Y=g(X)$, $F_Y(y)=\int_{g(x)\le y} f(x)\,dx$

$g(x)$分段单调，各段有 $x=h_j(Y)$, $f_Y(y)=\sum f(h_j(Y))\,|h_j'(Y)|\,I_j$

三、多维随机变量及其分布

joint cdf $\quad F(\vec{x})=P(\vec{X}\le\vec{x})$

pmf $\quad P(\cap X_i=x_{ij_i})=p_{j_1\cdots j_n}$

pdf $\quad f(\vec{x})\quad F(\vec{x})=\int_{(-\infty)^n}^{\vec{x}} f(\vec{u})\,d\vec{u}$

边缘分布 $\quad P(X_i=x_{ij_i})=\sum\limits_{j_1\cdots j_{i-1},j_{i+1}\cdots j_n} p_{j_1\cdots j_n}$

$\vec{X}\to\vec{U}+\vec{V}\quad F_{\vec{U}}(\vec{u})=P(\vec{U}\le\vec{u})$

$f_{\vec{U}}(\vec{u})=\int_{R^{n-k}} f(\vec{x})\,d\vec{v}$

条件分布 $\quad \vec{X}\to\vec{U}+\vec{V}\quad \vec{U}\sim f_{\vec{U}}(\vec{u})\quad \vec{V}\sim f_{\vec{V}}(\vec{v})$, $f_{\vec{U}|\vec{V}}(\vec{u}|\vec{v})=\frac{f(\vec{x})}{f_{\vec{V}}(\vec{v})}$

独立
$F(\vec{x})=F_1(x_1)\cdots F_n(x_n)$
$f(\vec{x})=f_1(x_1)\cdots f_n(x_n)$
$P(\cap X_i=x_{ij_i})=\prod P(X_i=x_{ij_i})$
$\Big\rangle\ x_1,\dots,x_n$相互独立

函数 $\quad \vec{Z}=\vec{g}(\vec{X})\quad F_{\vec{Z}}(\vec{z})=\int\!\!\int_{\vec{g}(\vec{x})\le\vec{z}} f(\vec{x})\,d\vec{x}$

$\vec{X}=\vec{\varphi}(\vec{Z})\quad F_{\vec{Z}}(\vec{z})=\int\!\!\int_{\vec{u}\le\vec{z}} f(\vec{\varphi}(\vec{u}))\left|\frac{\partial\vec{\varphi}}{\partial\vec{u}}\right| d\vec{u}$

$f_{\vec{Z}}(\vec{z})=f(\vec{\varphi}(\vec{u}))\left|\frac{\partial\vec{\varphi}}{\partial\vec{u}}\right|_{\vec{u}=\vec{z}}$

$X \sim F_1(x),\ f_1(x)$ $Y \sim F_2(y),\ f_2(y)$，X, Y 独立

$f_{X+Y}(z) = f_1 * f_2(z)$

$f_{\cancel{XY}}(z) = \int_{-\infty}^{+\infty} |v| f_1(vz) f_2(v)\, dv$

$F_{max}(z) = F_1(z) F_2(z)$

$F_{min}(z) = 1-(1-F_1(z))(1-F_2(z))$

四、随机变量的数字特征和极限定理

数学期望 $E(X) = \begin{cases} \sum x_k p_k \\ \int x f_X(x)\, dx \end{cases}$

(1) 线性：$E(X_1+\dots+X_n) = \sum_1^n X_i$

(2) $X_1, \dots, X_n$ 相互独立，$E(\prod X_i) = E(X_1) \cdots E(X_n)$

(3) $\vec{Y} = g(\vec{X})$，$\vec{X} \sim f_{\vec{X}}(\vec{x})$，$E(\vec{Y}) = \begin{cases} \int g(\vec{x}) f_{\vec{X}}(\vec{x})\, d\vec{x} \\ \sum g(\vec{x}) P(\vec{X} = \vec{x}) \end{cases}$

(4) $X \geq Y$，$E(X) \geq E(Y)$

条件期望 $E(Y|X=x) = \int y f_{Y|X}(y|X=x)\, dy$

全期望公式：$E[E(Y|X=x)] = E(Y)$

$E(\cdot|X=x)$ 是期望函数

中位数 m　$P(X \geq m) \geq 0.5$，$P(X \leq m) \geq 0.5$

众数 m_d．$P(X = m_d) = \max p$，$f(m_d) = \max f(x)$

p 分位数 Q_p　$P(X \leq Q_p) \geq p$，$P(X \geq Q_p) \leq 1-p$

内四分位距 $IQR(X) = Q_{0.75} - Q_{0.25}$

方差和标准差 $\sigma^2 = Var(X) = E(X-\mu)^2$，$\sigma = \sqrt{Var(X)}$

(1) $Var(X) = E(X^2) - \mu^2$

(2) $Var(c) = 0$

13) $Var(cX) = c^2 Var(X)$, $Var(X+c) = Var(X)$

14) $Var(\Sigma X_i) = \Sigma Var(X_i)$，如果 $X_1, \ldots, X_n$ 两两独立

15) $Var(X) \geq 0$，$P(X=E(X))$ 等号成立

$Var(X) \leq E(X-c)^2$，$c=E(X)$ 等号成立

$Y = \frac{X-\mu}{\sigma}$ 为标准化，无量纲

马尔可夫不等式：$Y \geq 0$，$\forall \varepsilon > 0$，$P(Y \geq \varepsilon) \leq \frac{E(Y)}{\varepsilon}$

切比雪夫不等式：$P(|X-\mu| \leq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$

矩：k阶原点矩 $\alpha_k = E(X^k)$

k阶中心矩 $\mu_k = E(X-\mu)^k$ $\beta_1 = \frac{\mu_3}{\sigma^3}$ $\beta_2 = \frac{\mu_4}{\sigma^4}(-3)$

矩母函数MGF $M_X(s) = E(e^{sX}) = \sum_0^{\infty} \frac{s^k}{k!} E(X^k)$

$$E(X^k) = \frac{d^k}{ds^k} M_X(s) \Big|_{s=0}$$

若 X, Y $\forall s \in [-c, c]$ 的矩母函数有限且相等，

那么 $F_X(t) = F_Y(t)$

$M_{X_1+\cdots+X_n}(s) = M_{X_1}(s) \cdots M_{X_n}(s)$，如果 $X_1, \ldots, X_n$ 独立

协方差和相关系数 $Cov(X, Y) = E[(X-E(X))(Y-E(Y))]$

$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

(1) $Cov(X, Y) = Cov(Y, X)$

(2) $Cov(X, Y) = E(XY) - E(X)E(Y)$

(3) $Cov(aX+bY, cX+dY) = (a, b)\begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix}$

(4) $Cov(X, Y) = 0$，如果 X, Y 独立

柯西-施瓦茨不等式：$[Cov(X, Y)]^2 \leq Var(X) Var(Y)$，$X$ 与 Y 有严格线性关系时等号成立

相关系数 $\rho_{X,Y} = Cov\left(\frac{X-E(X)}{\sqrt{Var(X)}}, \frac{Y-E(Y)}{\sqrt{Var(Y)}}\right) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

$\rho_{X,Y} = 0$，如果 X, Y 独立

熵 $H(X)=\begin{cases}-\sum p_k \log_2(p_k) \\ -\int f_X(x)\ln f_X(x)dx\end{cases}$

$H(X+c)=H(X)$

离散型(1) $H(X)\geq 0$; (2) $H(X)\leq \log_2(n)$，当离散均匀分布时等号成立

熵最大化优化问题：

$\max\limits_{-\infty<X<\infty} H(X)$, s.t. $\mu,\sigma^2 \longrightarrow X\sim N(\mu,\sigma^2)$

$\max\limits_{-\infty<X<\infty} H(X)$, s.t. $\frac{1}{\lambda} \longrightarrow X\sim Exp(\lambda)$

$\max\limits_{a\leq X\leq b} H(X) \longrightarrow X\sim \cancel{\frac{1}{b-a}I_{[a,b]}}\ U[a,b]$

依概率收敛：$X_1,\dots,X_n$ ~~i.i.d~~ $\forall\varepsilon>0$, $\lim\limits_{n\to\infty} P(|X_n-X|\geq\varepsilon)=0$,

称 $X_n\to X$, in P 或 $X_n\xrightarrow{P}X$

大数定律：$X_1,\dots,X_n$ i.i.d, μ,σ^2, 记 $S_n=\sum X_i$, $\forall\varepsilon>0$, $\lim\limits_{n\to\infty}P(|\frac{S_n}{n}-\mu|\geq\varepsilon)=0$

伯努利大数定律：$\{X_i\}$ i.i.d 0-1分布(p)，$\frac{1}{n}\sum X_k\xrightarrow{P}p$

依分布收敛：$X_i\sim F_i(x)$，$\lim\limits_{n\to\infty}F_n(x)=F(x)$，称$\{F_n(x)\}$弱收敛于$F(x)$，$X_n\xrightarrow{L}X$

$X_n\xrightarrow{P}X\Rightarrow X_n\xrightarrow{L}X$，$X_n\xrightarrow{L}c\Rightarrow X_n\xrightarrow{P}c$

林德伯格-莱维中心极限定理：$X_1,\dots,X_n$ i.i.d, μ,σ^2, $S_n=\sum X_i$, $\forall x\in R$

$\lim\limits_{n\to\infty}P(\frac{\sqrt{n}(S_n/n-\mu)}{\sigma}\leq x)=\Phi(x)$，即 $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\xrightarrow{L}N(0,1)$.

一般 $n>30$

棣莫夫-拉普拉斯中心极限定理 $X_1,\dots,X_n$ i.i.d $\sim B(1,p)$, $S_n=\sum X_i$

$\forall x\in R$, $\lim\limits_{n\to\infty}P(\frac{S_n-np}{\sqrt{np(1-p)}}\leq x)=\Phi(x)$；逼近二项分布两端需修正

第五章 统计学基本概念

χ^2分布：$N(0,1) \longrightarrow (X_1,\dots,X_n)$　$X = X_1^2+\dots+X_n^2 \sim \chi_n^2$

(1) $E(X)=n$，$Var(X)=2n$

(2) $X\sim\chi_m^2$，$Y\sim\chi_n^2$，X,Y独立，$X+Y\sim\chi_{m+n}^2$

t分布：$X\sim N(0,1)$，$Y\sim\chi_n^2$，X,Y独立　$T=\frac{X}{\sqrt{Y/n}}\sim t_n$

(1) $n=1$，T为柯西分布

(2) $n\geq 2$，$E(T)=0$，$n\geq 3$，$Var(T)=\frac{n}{n-2}$

(3) $n\to\infty$，$t_n \xrightarrow{L} \Phi(x)$

F分布：$X\sim\chi_m^2$，$Y\sim\chi_n^2$，X,Y独立，$F=\frac{X/m}{Y/n}\sim F_{m,n}$

(1) $1/F\sim F_{n,m}$

(2) $T\sim t_n$，$T^2\sim F_{1,n}$

(3) $F_{m,n}(1-\alpha)=1/F_{n,m}(\alpha)$

$X_1,\dots,X_n$ i.i.d $\sim N(\mu,\sigma^2)$

(1) 独立正态随机变量的线性组合服从正态分布

$\sum c_k X_k \sim N(\mu\sum c_k, \sigma^2\sum c_k^2)$，$c_k$不全为0

(2) $S^2=\frac{1}{n-1}\sum(X_i-\bar{X})^2$，$\frac{(n-1)S^2}{\sigma^2}\sim\chi_{n-1}^2$

(3) $\bar{X}$与S^2相互独立

(4) $\frac{\sqrt{n}(\bar{X}-\mu)}{S}\sim t_{n-1}$

$X_1,\dots,X_m$ i.i.d $\sim N(\mu_1,\sigma^2)$，$Y_1,\dots,Y_n$ i.i.d $\sim N(\mu_2,\sigma^2)$，$X_1,\dots,X_m$与$Y_1,\dots,Y_n$独立，$T=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_T}\sqrt{\frac{mn}{m+n}}\sim t_{m+n-2}$，

其中 $(n+m-2)S_T^2=(m-1)S_X^2+(n-1)S_Y^2$

$X_1,\dots,X_m$ i.i.d $\sim N(\mu_1,\sigma_1^2)$，$Y_1,\dots,Y_n$ i.i.d $\sim N(\mu_2,\sigma_2^2)$，$X_1,\dots,X_m$与$Y_1,\dots,Y_n$独立，$F=\frac{S_X^2}{S_Y^2}\frac{\sigma_2^2}{\sigma_1^2}\sim F_{m-1,n-1}$

$X_1,\dots,X_n$ i.i.d $\sim Exp(\lambda)$，$2\lambda n\bar{X}\sim\chi_{2n}^2$

六、参数点估计

点估计 $\theta \to \hat{\theta}$

矩估计法 $\hat{\alpha}_j = a_j$，$\hat{\mu}_j = m_j$

最大似然估计 $(X_1, \dots, X_n) \sim f(\vec{x};\theta) = L(\theta;\vec{x})$

$L(\hat{\theta}_L;\vec{x}) = \max\limits_{\theta\in\Theta} L(\theta;\vec{x})$

$l = \ln L$，比较驻点、边界

优良性准则

无偏性
- $\theta \to \hat{\theta}$，$E_\theta(\hat{\theta}) - \theta$ 为 $\hat{\theta}$ 偏差，$\forall \theta \in \Theta$，若 $E_\theta(\hat{\theta}) = \theta$，称 $\hat{\theta}$ 为 θ 一个无偏估计。无偏估计不一定存在，也不一定唯一。$\hat{\theta}_M$，$\hat{\theta}_L$ 一般不是 θ 的无偏估计，需修正。
- 若 $\lim\limits_{n\to\infty} E_\theta(\hat{\theta}) = \theta$，$\forall \theta \in \Theta$，称 $\hat{\theta}$ 是 θ 的一个渐近无偏估计，$\hat{\theta}_M$ 是 θ 的渐进无偏估计。

有效性
- $MSE_\theta(\hat{\theta}) = E_\theta(\hat{\theta} - \theta)^2 = Var_\theta(\hat{\theta})$（无偏）
- $\theta \to \hat{\theta}_1, \hat{\theta}_2$ 无偏，若 $Var_\theta(\hat{\theta}_1) \le Var_\theta(\hat{\theta}_2)$，$\theta \in \Theta$，且至少一个 θ 使等号成立，称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效
- $\theta \to \{\hat{\theta}\}$ 无偏，若 $Var_\theta(\hat{\theta}_1) = \min Var_\theta(\hat{\theta})$，$\theta \in \Theta$，称 $\hat{\theta}_1$ 是 θ 最小方差无偏估计

相合性 $\theta \to \hat{\theta}$，若 $n \to \infty$，$\hat{\theta} \xrightarrow{P} \theta$，称 $\hat{\theta}$ 是 θ 弱相合估计量。$\hat{\theta}_M$，$\hat{\theta}_L$（一般）是 θ 相合估计

渐进正态性 $\theta \to \hat{\theta}$，$Var_\theta(\hat{\theta}) = \sigma_n^2(\theta)$，若 $n \to \infty$，$\dfrac{\hat{\theta} - \theta}{\sigma_n(\theta)} \xrightarrow{L} N(0,1)$，称 $\hat{\theta}$ 有渐进正态性。一般 $\hat{\theta}_M$、$\hat{\theta}_L$ 有渐进正态性。

七、区间估计

置信区间和置信限

$F(x;\theta) \to (X_1,\dots,X_n)$，$\alpha\in(0,1)$，若 $P_\theta(\theta_1 \le \theta \le \theta_2) = 1-\alpha$，$\forall \theta \in \Theta$，称 $[\theta_1,\theta_2]$ 为 θ 置信系数 $1-\alpha$ 的置信区间。$P_\theta(\theta_1 \le \theta \le \theta_2) \ge \beta$，$\beta$ 为置信水平。若 $P_\theta(\bar{\theta} \ge \theta) = P_\theta(\underline{\theta} \le \theta) = 1-\alpha$，$\bar{\theta}$（$\underline{\theta}$）为置信上（下）限

枢轴变量法

正态总体均值 σ^2 已知，$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$；$\sigma^2$ 未知，$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$

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

两～ 均值差 σ_1^2,σ_2^2 已知 $\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \sim N(0,1)$

~~方差比~~ $\sigma_1^2=\sigma_2^2$ 未知 $\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_T}\sqrt{\frac{mn}{m+n}} \sim t_{m+n-2}$

方差比 $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{m-1,n-1}$

大样本方法

0-1总体概率 ~~0-1分布 二项总体~~ $n\to\infty$（一般大于100）$\frac{Y_n-np}{\sqrt{np(1-p)}} \sim N(0,1)$

当 $n\hat{p}$，$n(1-\hat{p})$ 大于10，可忽略 $u^2_{\alpha/2}/n$

当 $q = p\times100\%$，以 q 代 p，以 $100-q$ 代 $1-p$

一般总体均值 $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$，$n\to\infty$

自助法

$F \to (x_1,\dots,x_n) \xrightarrow{B\text{ 一般}1000} (x_1^*,\dots,x_n^*)$，$T=\hat{\theta}-\theta \sim \hat{\theta}^*-\hat{\theta}$

$\hat{\theta} \xrightarrow{\text{自助分布}} \hat{\theta}^*$

自助 t 置信区间：$\theta \to \hat{\theta}$（渐近无偏），$T=\frac{\hat{\theta}-\theta}{se(\hat{\theta})}$，在总体未知时以自助法得 $\hat{se}(\hat{\theta}) = \sqrt{\frac{1}{B-1}\Sigma(\hat{\theta}_i^*-\bar{\hat{\theta}^*})^2}$，于是 T 的分布以 $T^* = \frac{\hat{\theta}^*-\hat{\theta}}{\hat{se}(\hat{\theta})}$ 近似，~~它是 t 分布（$\hat{se}(\hat{\theta})$ 亦需要自助）~~ 再利用自助法获得 ~~T* 分布~~ $\hat{se}(\hat{\theta}^*)$，得到 T^* 的分布

八、假设检验

原假设与备择假设、检验函数、功效函数、两类错误、显著性检验

~~一样本~~正态总体参数的检验 { 均值（σ^2已知、未知）、均值差（成组、成对）、方差、方差比

比例p的检验 { 精确方法、大样本方法

似然比检验

$H_0: \theta \in \Theta_0 \leftrightarrow H_1: \theta \in \Theta / \Theta_0$，$\vec{x} \sim f(\vec{x}, \theta)$

$$LR(\vec{x}) = \sup_{\theta \in \Theta} f(\vec{x};\theta) \Big/ \sup_{\theta \in \Theta_0} f(\vec{x};\theta)$$

$$2\ln LR(\vec{x}) \xrightarrow{\mathcal{L}} \chi^2_{k-s}$$

p值

p值$=P$(得到$T(\vec{x})$或更极端值$|H_0$)，ψ：p值$<\alpha$时拒绝H_0

置信区间$\longleftrightarrow$假设检验

九、非参数假设检验

拟合优度检验

$H_0: P(X=a_i)=p_i$，$i=1,\dots,k$

$E=np_i$ $O=n_i$

$n\to\infty$，$Z=\sum \frac{(O-E)^2}{E} \xrightarrow{\mathcal{L}} \chi^2_{k-1}$

拟合优度$p(Z_0)=P(Z \geq Z_0)$

$H_0': P(X=a_i)=p_i(\theta_1,\dots,\theta_r)$，$i=1,\dots,k$

$\hat{E}=n\hat{p}_i$

$n\to\infty$，$Z=\sum \frac{(O-\hat{E})^2}{\hat{E}} \xrightarrow{\mathcal{L}} \chi^2_{k-1-r}$

$H_0'': X \sim F_\theta(x)$，$x \in R$

$Y=a_i$，$x_{i-1} \leq X < x_i$，转化为H_0'

列联表检验

$$n\to\infty \quad Z=\sum\sum \frac{(n_{ij}-n_{i\cdot}n_{\cdot j}/n)^2}{n_{i\cdot}n_{\cdot j}/n} \xrightarrow{\mathcal{L}} \chi^2_{(a-1)(b-1)}$$

经验法则：$n \geq 50$，$n\hat{p}_i > 5$