§6 平均、分散、標準偏差

6.1 平均値(期待値)

(i) Xの平均値：$m=E(X)=\begin{cases}\sum_i x_i p_i & \text{(離散的)}\\ \int_{-\infty}^{\infty} x f(x)dx & \text{(連続的)}\end{cases}$

(ii) $E(g(X))=\begin{cases}\sum_i g(x_i)p_i & \text{(離散的)}\\ \int_{-\infty}^{\infty} g(x)f(x)dx & \text{(連続的)}\end{cases}$

(iii) $E(a_1X_1+\cdots+a_kX_k+b)=a_1E(X_1)+\cdots+a_kE(X_k)+b$ ($a_1,\cdots a_k,b$:定数)

(iv) $X_1,X_2,\cdots X_k$が独立 $\Rightarrow E(X_1X_2\cdots X_k)=E(X_1)E(X_2)\cdots E(X_k)$

6.2 分散，標準偏差

(i) Xの分散：$\sigma^2=V(X)=E[(X-E(X))^2]=E(X^2)-E^2(X)$

(ii) Xの標準偏差：$\sigma=\sqrt{V(X)}=\sqrt{E(X^2)-E^2(X)}$ (≥ 0)

(iii) $V(aX+b)=a^2V(X)$

(iv) $X_1,\cdots X_k$が独立 $\Rightarrow V(a_1X_1+\cdots a_kX_k+b)=a_1^2V(X_1)+\cdots+a_k^2V(X_k)$

6.3 共分散，相関係数

(i) 共分散 $Cov(aX,bY)=ab\,Cov(X,Y)$ $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

$Cov(X_1,X_2)=\sigma_{X_1X_2}=E[(X_1-E(X_1))(X_2-E(X_2))]=E(X_1X_2)-E(X_1)E(X_2)$

(ii) 相関係数

$\rho=\rho_{X_1X_2}=\frac{\sigma_{X_1X_2}}{\sigma_{X_1}\sigma_{X_2}}=\frac{Cov(X_1,X_2)}{\sigma_{X_1}\sigma_{X_2}}$

(iii) $V(X_1+X_2)=V(X_1)+V(X_2)+2Cov(X_1,X_2)$

(iv) X_1,X_2が独立 $\Rightarrow Cov(X_1,X_2)=0$

$V(X_1+X_2)=V(X_1)+V(X_2)$

(v) $X_1,X_2,\cdots X_n$に対して、$Var\,X_i=\sigma_{ii}$

$Cov(X_i,X_j)=\sigma_{ij}=\sigma_{ji}$ $\rho_{ij}=\rho_{ji}=\rho_{X_i,X_j}$, $\rho_{ii}=1$とし、

$(\sigma_{ij})=\begin{bmatrix}\sigma_{11}&\sigma_{12}&\cdots&\sigma_{1n}\\ \sigma_{21}&\sigma_{22}&\cdots&\sigma_{2n}\\ \cdots&\cdots&\cdots&\cdots\\ \sigma_{n1}&\sigma_{n2}&\cdots&\sigma_{nn}\end{bmatrix}$ $(\rho_{ij})=\begin{bmatrix}\rho_{11}&\rho_{12}&\cdots&\rho_{1n}\\ \rho_{21}&\rho_{22}&\cdots&\rho_{2n}\\ \cdots&\cdots&\cdots&\cdots\\ \rho_{n1}&\rho_{n2}&\cdots&\rho_{nn}\end{bmatrix}$ 共分散行列 相関行列

確率変数X，期待値$E(X)=\mu$，分散$V(X)=\sigma^2$

Chebyshev不等式：$P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}$ (チェビシェフの不等式)

$E(X^k)$ X的k阶原点矩　$E(X^kY^l)$ X和Y的k+l阶混合矩

$E\{[X-E(X)]^k\}$ X的k阶中心矩　$E\{[X-E(X)]^k[Y-E(Y)]^l\}$ X和Y的k+l阶混合中心矩

$C_{11}=E\{[X_1-E(X_1)]^2\}$

$C_{12}=E\{[X_1-E(X_1)][X_2-E(X_2)]\}$

$C_{21}=E\{[X_2-E(X_2)][X_1-E(X_1)]\}$

$C_{22}=E\{[X_2-E(X_2)]^2\}$

$\begin{pmatrix}C_{11}&C_{12}\\C_{21}&C_{22}\end{pmatrix}$ 共分散行列

§7 大数法則と中心極限定理

7.1 大数の(弱)法則

確率変数$X_1,X_2,\cdots,X_n$が独立，$E(X_i)=\mu$，$V(X_i)=\sigma^2$，$\bar{X}=\sum_{i=1}^n X_i/n$，$\varepsilon>0$のとき

$\lim_{n\to\infty}P(|\bar{X}-\mu|\geq\varepsilon)=0$

⇒ Bernoulli 大数法則　f_A是n次独立重复试验中事件A发生的次数，P是事件A在每次试验中发生的概率，对任意正数$\varepsilon>0$，有

$\lim_{n\to\infty}P\{|\frac{f_A}{n}-p|<\varepsilon\}=1$

7.2 中心極限定理

確率変数$\{X_i\}$が同一分布に従い，互いに独立で，$E(X_i)=\mu$，$V(X_i)=\sigma^2$ $(i=1,2\cdots)$のとき

$\lim_{n\to\infty}P\left(\frac{X_1+X_2+\cdots+X_n-n\mu}{\sqrt{n\sigma^2}}\leq x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt=N(0,1)$ $\Rightarrow\bar{X}\sim N(\mu,\sigma^2/n)$

7.3 Lyapunov 定理

確率変数$\{X_i\}$が互いに独立で，$E(X_k)=\mu_k$，$V(X_k)=\sigma_k^2>0$

$\frac{\sum_{k=1}^n X_k-\sum_{k=1}^n\mu_k}{\sqrt{\sum_{k=1}^n\sigma_k^2}}\sim N(0,1)\Rightarrow\sum_{k=1}^n X_k\sim N\left(\sum_{k=1}^n\mu_k,\sum_{k=1}^n\sigma_k^2\right)$

7.4 De Moivre-Laplace 定理

確率変数$\{\eta_n\}$が二项分布$B(n,p)$に従い

$\lim_{n\to\infty}P\left\{\frac{\eta_n-np}{\sqrt{np(1-p)}}\leq x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt=\Phi(x)$

例 $X_k=0$ $p=0.05$
$X_k=1$ $p=0.8$
$X_k=2$ $p=0.15$

(1) $P(X>450)=?$　$X=\sum_{k=0}^{400}X_k$

X_k	0	1	2
P_k	0.05	0.8	0.15

$E(X_k)=1.1$　$D(X_k)=0.19$

$P\{X>450\}=P\left\{\frac{X-400\times1.1}{\sqrt{400}\sqrt{0.19}}>\frac{450-400\times1.1}{\sqrt{400}\sqrt{0.19}}\right\}$

$=1-P\left\{\frac{X-400\times1.1}{\sqrt{400}\sqrt{0.19}}\leq1.147\right\}\approx1-\Phi(1.147)=0.1251$

(2) $Y\sim b(400,0.8)$

$P(Y\leq340)$

$=P\left(\frac{Y-400\times0.8}{\sqrt{400\times0.8\times0.2}}\leq\frac{340-400\times0.8}{\sqrt{400\times0.8\times0.2}}\right)$

$\approx\Phi(2.5)=0.9938$

分布：

I. ガンマ分布$\Gamma(\lambda,\alpha)$

(i) 確率密度関数：$f(x)=\begin{cases}\frac{\alpha^\lambda}{\Gamma(\lambda)}x^{\lambda-1}e^{-\alpha x}&(0<x)\\0&(x\leq0)\end{cases}$ $(\alpha>0,\lambda>0)$　(ii) 平均値$=\frac{\lambda}{\alpha}$ 分散$=\frac{\lambda}{\alpha^2}$

II. χ^2分布$\chi^2(n)$

(i) 確率密度関数：$f(x)=\begin{cases}\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2}&(0<x)\\0&(x\leq0)\end{cases}$ (自由度n)　(ii) 平均値$=n$ 分散$=2n$

III. t分布(スチューデント分布) $t(n)$

(i) 確率密度関数：$f(x)=\frac{1}{\sqrt{n}}\frac{1}{B(\frac{1}{2},\frac{n}{2})}(1+\frac{x^2}{n})^{-(n+1)/2}$ $(-\infty<x<\infty)$ (自由度n)　(ii) 平均値=0 $(n>1)$ 分散$n/(n-2)$ $(n>2)$

IV. F分布(スネデッカー分布) $F(m,n)$

(i) 確率密度関数：$f(x)=\begin{cases}\frac{m^{m/2}n^{n/2}}{B(\frac{m}{2},\frac{n}{2})}\frac{x^{m/2-1}}{(mx+n)^{(m+n)/2}}&(0<x)\\0&(x\leq0)\end{cases}$　(ii) 平均値$=\frac{n}{n-2}$ $(n>2)$ 分散$=\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $(n>4)$

§8 参数估计

总体X的分布函数$F(x;\theta)$的形式已知，θ为待估参数．$X_1,X_2\cdots X_n$为X的一个样本，$x_1,x_2,x_3\cdots$为相应样本值．$\hat{\theta}(x_1,x_2,\cdots x_n)$为θ估计值

(I) 矩估计法

$\begin{cases}\mu_1=E(X)=\mu_1(\theta_1,\theta_2,\cdots,\theta_k)\\ \mu_2=E(X^2)=\mu_2(\theta_1,\theta_2,\cdots,\theta_k)\\ \vdots\\ \mu_k=E(X^k)=\mu_k(\theta_1,\theta_2,\cdots,\theta_k)\end{cases}$ 可以解出 $\begin{cases}\theta_1=\theta_1(\mu_1,\mu_2,\cdots\mu_k)\\ \theta_2=\theta_2(\mu_1,\mu_2,\cdots\mu_k)\\ \vdots\\ \theta_k=\theta_k(\mu_1,\mu_2,\cdots\mu_k)\end{cases}$ 用$A_l=\frac{1}{n}\sum_{i=1}^n X_i^l$替换对应$\mu_l$即可

(II) 最大似然估计法

在参数为θ的情况下，事件$\{X_1=x_1,X_2=x_2,\cdots X_n=x_n\}$发生的概率为$L(\theta)=L(x_1,x_2,\cdots x_n;\theta)=\prod_{i=1}^n p(x_i;\theta)$

当$\theta=\theta_0$使$L(\theta)$取到最大值时

将θ_0这一值称为最大似然估计值．常可从方程$\frac{d}{d\theta}L(\theta)=0$中解得

对数似然方程 $\frac{d}{d\theta}\ln L(\theta)=0$

多参数$\theta_1,\theta_2,\cdots\theta_k$情况：

$\frac{\partial}{\partial\theta_i}L=0,\ i=1,2,\cdots,k$　$\frac{\partial}{\partial\theta_i}\ln L=0,\ i=1,2,\cdots,k$ 对数似然方程组

最大似然估计的不变性

设θ的函数$u=u(\theta)$　$L(x_1,x_2,\cdots x_n;\theta(\hat{u}))=\max L(x_1,x_2,\cdots x_n;\theta(u))$

8.3 估计量的评选标准

(1)无偏性 估计量$\hat{\theta}$的数学期望存在，且$E(\hat{\theta})=\theta$ (θ为待估参数) 则$\hat{\theta}$为θ的无偏估计量

总体X均值为μ，方差为$\sigma^2>0$未知

样本$\bar{X}$取为总体均值μ的无偏估计

样本方差$S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$为总体方差$\sigma^2$的无偏估计

k阶样本矩$A_k=\frac{1}{n}\sum_{i=1}^n X_i^k$是k阶总体矩$\mu_k$的无偏估计量

(2)有效性

如果在样本容量n相同的情况下，$\hat{\theta}_1$较$\hat{\theta}_2$更密集的在真值θ附近，我们就认为$\hat{\theta}_1$较$\hat{\theta}_2$理想

若$\hat{\theta}_1$与$\hat{\theta}_2$都是θ的无偏估计量，$D(\hat{\theta}_1)\leq D(\hat{\theta}_2)$，则称$\hat{\theta}_1$较$\hat{\theta}_2$有效．

(3)相合性

若对于任意$\theta\in\Theta$，当$n\to\infty$时$\hat{\theta}(X_1,X_2,\cdots,X_n)$依概率收敛于θ，则称$\hat{\theta}$为θ的相合估计量

8.4 区间估计

置信区间：两个统计量$\underline{\theta}=\underline{\theta}(X_1,X_2,\cdots X_n)$，$\bar{\theta}=\bar{\theta}(X_1,X_2,\cdots X_n)$ $\underline{\theta}<\bar{\theta}$　$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

$P\{\underline{\theta}(X_1,X_2,\cdots X_n)<\theta<\bar{\theta}(X_1,X_2,\cdots,X_n)\}\geq1-\alpha$

$(\underline{\theta},\bar{\theta})$是θ的置信水平为$1-\alpha$的置信区间　函数$W=W(X_1,X_2,\cdots X_n;\theta)$的分布

$\underline{\theta}$ 置信下限　$1-\alpha$称为置信水平　不依赖于θ以及其他未知参数，称该函数

$\bar{\theta}$ 置信上限　为枢轴量．例 正态分布转化为$N(0,1)$