概率论与数理统计(经管类)公式

一、随机事件和概率

1、随机事件及其概率

运算律名称	表达式
交换律	$A+B=B+A \qquad AB=BA$
结合律	$(A+B)+C=A+(B+C)=A+B+C \qquad (AB)C=A(BC)=ABC$
分配律	$A(B\pm C)=AB\pm AC \qquad A+(BC)=(A+B)(A+C)$
德摩根律	$\overline{A+B}=\overline{A}\,\overline{B} \qquad \overline{AB}=\overline{A}+\overline{B}$

2、概率的定义及其计算

公式名称	公式表达式
求逆公式	$P(\overline{A})=1-P(A)$
加法公式	$P(A+B)=P(A)+P(B)-P(AB)$
条件概率公式	$P(B\|A)=\dfrac{P(AB)}{P(A)}$
乘法公式	$P(AB)=P(A)P(B\|A) \qquad P(AB)=P(B)P(A\|B)$
全概率公式	$P(B)=\sum_{i=1}^{n}P(A_i)P(B\|A_i)$
贝叶斯公式 （逆概率公式）	$P(A_j\|B)=\dfrac{P(A_j)P(B\|A_j)}{\sum_{i=1}^{\infty}P(A_j)P(B\|A_i)}$
伯努力概型公式	$P_n(k)=C_n^k p^k(1-p)^{n-k},k=0,1,\cdots n$
两件事件相互独立相应公式	$P(AB)=P(A)P(B)$； $P(B\|A)=P(B)$； $P(B\|A)=P(B\|\overline{A})$； $P(B\|A)+P(\overline{B}\|\overline{A})=1$； $P(\overline{B}\|A)+P(B\|\overline{A})=1$

二、随机变量及其分布

1、分布函数性质

$$P(X\le b)=F(b) \qquad P(a<X\le b)=F(b)-F(a)$$

2、离散型随机变量

分布名称	分布律
0－1 分布 $B(1,p)$	$P(X=k)=p^k(1-p)^{1-k},\quad k=0,1$
二项分布 $B(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k},\quad k=0,1,\cdots,n$

泊松分布 $P(\lambda)$	$P(X=k)=e^{-\lambda}\dfrac{\lambda^k}{k!},\quad k=0,1,2,\cdots$
几何分布 $G(p)$	$P(X=k)=(1-p)^{k-1}p,\quad k=0,1,2,\cdots$
超几何分布 $H(N,M,n)$	$P(X=k)=\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n},k=l,l+1,\cdots,\min(n,M)$

3、连续型随机变量

分布名称	密度函数	分布函数
均匀分布 $U(a,b)$	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & 其他\end{cases}$	$F(x)=\begin{cases}0, x<a\\ \dfrac{x-a}{b-a}, a\le x<b\\ 1, x\ge b\end{cases}$
指数分布 $E(\lambda)$	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & 其他\end{cases}$	$F(x)=\begin{cases}0, & x<0\\ 1-e^{-\lambda x}, & x\ge 0\end{cases}$
正态分布 $N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\ \sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\quad -\infty<x<+\infty$	$F(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t$
标准正态分布 $N(0,1)$	$\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\quad -\infty<x<+\infty$	$F(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t$

三、多维随机变量及其分布

1、离散型二维随机变量边缘分布

$$p_{i\cdot}=P(X=x_i)=\sum_j P(X=x_i,Y=y_j)=\sum_j p_{ij}\qquad p_{\cdot j}=P(Y=y_j)=\sum_i P(X=x_i,Y=y_j)=\sum_i p_{ij}$$

2、离散型二维随机变量条件分布

$$p_{i|j}=P(X=x_i|Y=y_j)=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{P_{\cdot j}},i=1,2\cdots$$

$$p_{j|i}=P(Y=y_j|X=x_i)=\frac{P(X=x_i,Y=y_j)}{P(X=x_i)}=\frac{p_{ij}}{P_{i\cdot}},j=1,2\cdots$$

3、连续型二维随机变量（X，Y）的分布函数 $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dvdu$

4、连续型二维随机变量边缘分布函数与边缘密度函数

分布函数：$F_X(x)=\int_{-\infty}^{x}\int_{-\infty}^{+\infty}f(u,v)dvdu$　　密度函数：$f_X(x)=\int_{-\infty}^{+\infty}f(x,v)dv$

$F_Y(y)=\int_{-\infty}^{y}\int_{-\infty}^{+\infty}f(u,v)dudv$　　$f_Y(y)=\int_{-\infty}^{+\infty}f(u,y)du$

5、二维随机变量的条件分布

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)},-\infty<y<+\infty\qquad f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)},-\infty<x<+\infty$$

四、随机变量的数字特征

1、数学期望

离散型随机变量：$E(X)=\sum_{k=1}^{+\infty}x_k p_k$　　连续型随机变量：$E(X)=\int_{-\infty}^{+\infty}xf(x)dx$

2、数学期望的性质

(1) $E(C)=C$,C为常数　$E[E(X)]=E(X)$　$E(CX)=CE(X)$

(2) $E(X\pm Y)=E(X)\pm E(Y)$　$E(aX\pm b)=aE(X)\pm b$　$E(C_1X_1+\cdots C_nX_n)=C_1E(X_1)+\cdots C_nE(X_n)$

(3)若 XY 相互独立则：$E(XY)=E(X)E(Y)$

(4) $[E(XY)]^2\le E^2(X)E^2(Y)$

3、方差：$D(X)=E(X^2)-E^2(X)$

4、方差的性质

(1) $D(C)=0$　$D[D(X)]=0$　$D(aX\pm b)=a^2D(X)$　$D(X)<E(X-C)^2$

(2) $D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$　若 XY 相互独立则：$D(X\pm Y)=D(X)+D(Y)$

5、协方差：$Cov(X,Y)=E(XY)-E(X)E(Y)$　若 XY 相互独立则：$Cov(X,Y)=0$

6、相关系数：$\rho_{XY}=\rho(X,Y)=\dfrac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$　若 XY 相互独立则：$\rho_{XY}=0$ 即 XY 不相关

7、协方差和相关系数的性质

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

(2) $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$　$Cov(aX+c,bY+d)=abCov(X,Y)$

8、常见数学分布的期望和方差

分布	数学期望	方差
0-1 分布 $B(1,p)$	p	$p(1-p)$
二行分布 $B(n,p)$	np	$np(1-p)$
泊松分布 $P(\lambda)$	λ	λ
几何分布 $G(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
超几何分布 $H(N,M,n)$	$n\frac{M}{N}$	$n\frac{M}{N}(1-\frac{M}{N})\frac{N-m}{N-1}$
均匀分布 $U(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布 $N(\mu,\sigma^2)$	μ	σ^2
指数分布 $E(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

五、大数定律和中心极限定理

1、切比雪夫不等式

若 $E(X)=\mu, D(X)=\sigma^2$, 对于任意 $\xi>0$ 有 $P\{|X-E(X)|\ge\xi\}\le\frac{D(X)}{\xi^2}$ 或 $P\{|X-E(X)|<\xi\}\ge 1-\frac{D(X)}{\xi^2}$

2、大数定律：若 $X_1 \cdots X_n$ 相互独立且 $n\to\infty$ 时，$\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{D} \frac{1}{n}\sum_{i=1}^{n}E(X_i)$

(1)若 $X_1 \cdots X_n$ 相互独立，$E(X_i)=\mu_i, D(X_i)=\sigma_i^2$ 且 $\sigma_i^2\le M$ 则：$\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n}E(X_i), (n\to\infty)$

(2)若 $X_1 \cdots X_n$ 相互独立同分布，且 $E(X_i)=\mu_i$ 则当 $n\to\infty$ 时：$\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{P} \mu$

3、中心极限定理

(1)独立同分布的中心极限定理：均值为 μ，方差为 $\sigma^2>0$ 的独立同分布时，当 n 充分大时有：

$$Y_n=\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\xrightarrow{\sim} N(0,1)$$

(2)拉普拉斯定理：随机变量 $\eta_n(n=1,2\cdots)\sim B(n,p)$ 则对任意 x 有：

$$\lim_{x\to+\infty}P\{\frac{\eta_n-np}{\sqrt{np(1-p)}}\le x\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x)$$

(3)近似计算：$P(a\le\sum_{k=1}^{n}X_k\le b)=P(\frac{a-n\mu}{\sqrt{n}\sigma}\le\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\le\frac{b-n\mu}{\sqrt{n}\sigma})\approx\Phi(\frac{b-n\mu}{\sqrt{n}\sigma})-\Phi(\frac{a-n\mu}{\sqrt{n}\sigma})$

六、数理统计

1、总体和样本

总体 X 的分布函数 $F(x)$ 样本 $(X_1,X_2\cdots X_n)$ 的联合分布为 $F(x_1,x_2\cdots x_n)=\prod_{k=1}^{n}F(x_k)$

2、统计量

(1)样本平均值：$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ (2)样本方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i^2-n\overline{X}^2)$

(3)样本标准差：$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$ (4)样本 k 阶原点距：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k, k=1,2\cdots$

(5)样本 k 阶中心距：$B_k=M_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k, k=2,3\cdots$

(6)次序统计量：设样本 $(X_1,X_2\cdots X_n)$ 的观察值 $(x_1,x_2\cdots x_n)$，将 $x_1,x_2\cdots x_n$ 按照由小到大的次序重新排列，得到 $x_{(1)}\le x_{(2)}\le\cdots\le x_{(n)}$，记取值为 $x_{(i)}$ 的样本分量为 $X_{(i)}$，则称 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$ 为样本 $(X_1,X_2\cdots X_n)$ 的次序统计量。$X_{(1)}=\min(X_1,X_2\cdots X_n)$ 为最小次序统计量；$X_{(n)}=\max(X_1,X_2\cdots X_n)$ 为最大次序统计量。

3、三大抽样分布

(1) χ^2 分布：设随机变量 $X_1,X_2,\cdots X_n$ 相互独立，且都服从标准正态分布 $N(0,1)$，则随机变量 $\chi^2=X_1^2+X_2^2+\cdots X_n^2$ 所服从的分布称为自由度为 n 的 χ^2 分布，记为 $\chi^2\sim\chi^2(n)$

性质：① $E[\chi^2(n)]=n,D[\chi^2(n)]=2n$ ②设 $X\sim\chi^2(m),Y\sim\chi^2(n)$ 且相互独立，则 $X+Y\sim\chi^2(m+n)$

(2) t 分布：设随机变量 $X\sim N(0,1),Y\sim\chi^2(n)$，且 X 与 Y 独立，则随机变量：$T=\frac{X}{\sqrt{Y/n}}$ 所服从的分布称为自由度的 n 的 t 分布，记为 $T\sim t(n)$

性质：① $E[t(n)]=0,D[t(n)]=\frac{n}{n-2},(n>2)$ ② $\lim\limits_{n\to\infty}t(n)=N(0,1)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

(3) F 分布：设随机变量 $U\sim\chi^2(n_1),V\sim\chi^2(n_2)$，且 U 与 V 独立，则随机变量 $F(n_1,n_2)=\frac{U/n_1}{V/n_2}$ 所服从的分布称为自由度 (n_1,n_2) 的 F 分布，记为 $F\sim F(n_1,n_2)$

性质：设 $X\sim F(m,n)$，则 $\frac{1}{X}\sim F(n,m)$

七、参数估计

1、参数估计

(1) 定义：用 $\hat{\theta}(X_1,X_2,\cdots X_n)$ 估计总体参数 θ，称 $\hat{\theta}(X_1,X_2,\cdots X_n)$ 为 θ 的估计量，相应的 $\hat{\theta}(X_1,X_2,\cdots X_n)$ 为总体 θ 的估计值。

(2) 当总体是正态分布时，未知参数的矩估计值=未知参数的最大似然估计值

2、点估计中的矩估计法：（总体矩=样本矩）

离散型样本均值：$\overline{X}=E(X)=\frac{1}{n}\sum_{i=1}^{n}X_i$　连续型样本均值：$\overline{X}=E(X)=\int_{-\infty}^{+\infty}xf(x,\theta)dx$

离散型参数：$E(X^2)=\frac{1}{n}\sum_{i=1}^{n}X_i^2$

3、点估计中的最大似然估计

最大似然估计法：$X_1,X_2,\cdots X_n$ 取自 X 的样本，设 $X\sim f(x,\theta)[或P(X=X_i)=P(\theta)]$ 则可得到概率密度：

$$f(x_1,x_2,\cdots x_n,\theta)=\prod_{i=1}^{n}f(x_i,\theta)[或P(X=X_1,X_2,\cdots X_n=x_n)=\prod_{i=1}^{n}P(X=x_i)=\prod_{i=1}^{n}P_i(\theta)]$$

基本步骤：

①似然函数：$L(\theta)=\prod_{i=1}^{n}f(x_i,\theta)[或\prod_{i=1}^{n}P_i(\theta)]$

②取对数：$\ln L=\sum_{i=1}^{n}\ln f(X_i,\theta)$

③解方程：$\frac{\partial \ln L}{\partial \theta_1}=0,\cdots,\frac{\partial \ln L}{\partial \theta_k}=0$ 最后得：$\hat{\theta}_1=\hat{\theta}_1(x_1,x_2,\cdots x_n),\cdots,\hat{\theta}_k=\hat{\theta}_k(x_1,x_2,\cdots x_n)$