随机事件与概率

公式名称	公式表达式
德摩根公式	$\overline{A\bigcup B}=\overline{A}\bigcap\overline{B}$，$\overline{A\bigcap B}=\overline{A}\bigcup\overline{B}$
古典概型	$P(A)=\frac{m}{n}=\frac{A\text{包含的基本事件数}}{\text{基本事件总数}}$
几何概型	$P(A)=\frac{\mu(A)}{\mu(\Omega)}$，其中 μ 为几何度量（长度、面积、体积）
求逆公式	$P(\overline{A})=1-P(A)$
加法公式	P(A∪B)=P(A)+P(B)-P(AB) 当 P(AB)＝0 时，P(A∪B)=P(A)+P(B)
减法公式	P(A-B)=P(A)-P(AB)，$B\subset A$ 时 P(A-B)=P(A)-P(B)
条件概率公式	$P(B\|A)=\frac{P(AB)}{P(A)}$
乘法公式	$P(AB)=P(A)P(B\|A)\quad P(AB)=P(B)P(A\|B)$
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）	$P(A_j\|B)=\frac{P(A_j)P(B\|A_j)}{\sum_{i=1}^{\infty}P(A_j)P(B\|A_i)}$
两件事件 相互独立	$P(AB)=P(A)P(B)$；$P(B\|A)=P(B)$；$P(B\|A)=P(B\|\overline{A})$；

二、随机变量及其分布

1、分布函数性质

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

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

3、续型随机变量

分布名称	密度函数	分布函数
均匀分布 $U(a,b)$	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{[illegible]}\end{cases}$	$F(x)=\begin{cases}0, x<a\\ \dfrac{x-a}{b-a}, a\le x<b\\ 1, x\ge b\end{cases}$

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

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

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

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

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 $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dudv$

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

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

6、X、Y 相互独立 $\Leftrightarrow F(x,y)=F_X(x)F_Y(y)\Leftrightarrow f(x,y)=f_X(x)f_Y(y)$

四、随机变量的数字特征

1、数学期望

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

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) $E(X\pm Y)=E(X)\pm E(Y)$　$E(aX\pm b)=aE(X)\pm b$

(3)若 X、Y 相互独立则：$E(XY)=E(X)E(Y)$

3、方差：$\boldsymbol{D}(\boldsymbol{X})=\boldsymbol{E}[(\boldsymbol{X}-\boldsymbol{E}(\boldsymbol{X}))^2]=\boldsymbol{E}(\boldsymbol{X}^2)-\boldsymbol{E}^2(\boldsymbol{X})$

4、方差的性质

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

(2) $D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$

(3)若 X、Y 相互独立则：$D(X\pm Y)=D(X)+D(Y)$

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

6、相关系数：$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$，X、Y 相互独立时：$\rho_{XY}=0$ (X, Y 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、常见随机变量分布的期望和方差

分布	数学期望	方差
0-1 分布 $b(1,p)$	p	$p(1-p)$
二项分布 $b(n,p)$	np	$np(1-p)$
泊松分布 $P(\lambda)$	λ	λ
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、大数定律与中心极限定理

1、切比雪夫不等式

若 $E(X)=\mu, D(X)=\sigma^2$, 对于任意 $\varepsilon>0$ 有 $\boldsymbol{P}\{|\boldsymbol{X}-\boldsymbol{E}(\boldsymbol{X})|\geq\varepsilon\}\leq\frac{\boldsymbol{D}(\boldsymbol{X})}{\varepsilon^2}$

或 $\boldsymbol{P}\{|\boldsymbol{X}-\boldsymbol{E}(\boldsymbol{X})|<\varepsilon\}\geq 1-\frac{\boldsymbol{D}(\boldsymbol{X})}{\varepsilon^2}$

2、大数定律：若 $X_1\cdots X_n$ 相互独立， $E(X_i)=\mu_i, D(X_i)=\sigma_i^2$ 且 $\sigma_{\boldsymbol{i}}^2\leq \boldsymbol{C}$

$$\text{则：}\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\frac{1}{n}\sum_{i=1}^{n}E(X_i),(n\to\infty)\text{（切比雪夫）}$$

若 $X_1\cdots X_n$ 相互独立同分布，且 $\boldsymbol{E}(\boldsymbol{X_i})=\mu$，则：$\frac{1}{\boldsymbol{n}}\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}}\boldsymbol{X_i}\xrightarrow[\boldsymbol{n}\to\infty]{\boldsymbol{P}}\mu$（辛钦）

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

(2)拉普拉斯定理：随机变量 $\boldsymbol{X}\sim\boldsymbol{B}(\boldsymbol{n},\boldsymbol{p})$ 则对任意 x 有：

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

(3)近似计算：

$$P(a\leq\sum_{k=1}^{n}X_k\leq b)=P(\frac{a-n\mu}{\sqrt{n}\sigma}\leq\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\leq\frac{b-n\mu}{\sqrt{n}\sigma})\approx\Phi(\frac{b-n\mu}{\sqrt{n}\sigma})-\Phi(\frac{a-n\mu}{\sqrt{n}\sigma})$$

六、数理统计的基本概念

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

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

(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

$F(n_1,n_2)=\dfrac{U/n_1}{V/n_2}$ 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 $X\sim F(n_1,n_2)$，则 $\dfrac{1}{X}\sim F(n_2,n_1)$

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=未知参数的极大似然估计值

2. 点估计中的矩估计法：（总体矩=样本矩）

样本均值：$\overline{X}=E(X)=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$ 或 $\overline{X}=E(X)=\int_{-\infty}^{+\infty}xf(x,\theta)dx$

求法步骤：设总体 X 的分布中包含有未知参数 $\theta_1,\theta_2,\cdots,\theta_k$，它的前 k 阶原点矩 $\mu_i=E(X^i)(i=1,2,\cdots,k)$ 中包含了未知参数 $\theta_1,\theta_2,\cdots,\theta_k$，即 $\mu_i=g_i(\theta_1,\theta_2,\cdots,\theta_k)(i=1,2,\cdots,k)$。又设 $x_1,x_2,\cdots,x_n$ 为总体 X 的 n 个样本值，用样本矩 $A_i=\dfrac{1}{n}\sum\limits_{j=1}^{n}X_j^i(i=1,2,\cdots,k)$ 代替 μ_i，在所建立的方程组中解出的 k 个未知参数即为参数

$\theta_1,\theta_2,\cdots,\theta_k$ 的矩估计量 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$

3. 点估计中的极大似然估计

极大似然估计法：$X_1,X_2,\cdots X_n$ 取自 X 的样本，设 $\boldsymbol{X}\sim\boldsymbol{f}(\boldsymbol{x},\theta)$ 或 $X\sim P(x,\theta)$，

求法步骤：

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：$\ln L(\theta)=\sum_{i=1}^{n}\ln f(x_i,\theta)$ 或 $\ln L(\theta)=\sum_{i=1}^{n}\ln p_i(\theta)$

③解方程：$\dfrac{\partial\ln L}{\partial\theta_1}=0,\cdots,\dfrac{\partial\ln L}{\partial\theta_k}=0$，解得：$\begin{cases}\hat{\theta}_1=\hat{\theta}_1(x_1,x_2,\cdots,x_n)\\ \cdots\cdots\\ \hat{\theta}_k=\hat{\theta}_k(x_1,x_2,\cdots,x_n)\end{cases}$

4. 估计量的评价标准

估计量的评价标准	无偏性	设 $\hat{\theta}=\hat{\theta}(x_1,x_2,\cdots,x_n)$ 为未知参数 θ 的估计量。若 E（$\hat{\theta}$）=θ，则称 $\hat{\theta}$ 为 θ 的无偏估计量。
	有效性	设 $\hat{\theta}_1=\hat{\theta}_1(x_1,x_2,\cdots,x_n)$ 和 $\hat{\theta}_2=\hat{\theta}_2(x_1,x_2,\cdots,x_n)$ 是未知参数 θ 的两个无偏估计量。若 $D(\hat{\theta}_1)<D(\hat{\theta}_2)$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效。
	一致性	设 $\hat{\theta}_n$ 是 θ 的一串估计量，如果对于任意的正数 ε，都有 $\lim\limits_{n\to\infty}P(\lvert\hat{\theta}_n-\theta\rvert>\varepsilon)=0,$ 则称 $\hat{\theta}_n$ 为 θ 的一致估计量（或相合估计量）。

5. 单正态总体参数的置信区间

条件	估计参数	枢轴量	枢轴量分布	置信水平为 $1-\alpha$ 的置信区间

已知 σ^2	μ	$Z=\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$	$N(0,1)$	$\left(\bar{x}-z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}},\bar{x}+z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right)$
未知 σ^2	μ	$T=\dfrac{\bar{x}-\mu}{S/\sqrt{n}}$	$t(n-1)$	$\left(\bar{x}-t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}},\bar{x}+t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}}\right)$
已知 μ	σ^2	$\chi^2=\sum_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma}\right)^2$	$\chi^2(n)$	$\left(\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)},\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right)$
未知 μ	σ^2	$\chi^2=\dfrac{(n-1)S^2}{\sigma^2}$	$\chi^2(n-1)$	$\left(\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$

八、假设检验

1. 假设检验的基本概念

基本思想	假设检验的统计思想是小概率原理。 这里所说的小概率事件就是事件 $\{K\in R_\alpha\}$，其概率就是显著性水平 α，通常我们取 α=0.05，有时也取 0.01 或 0.10。	
基本步骤	1. 提出原假设 H_0；2. 选择统计量 K；3. 对于 α 查表找分位数 λ； 4. 由样本值 $x_1,x_2,\cdots,x_n$ 计算统计量之值 K；将 $\hat{K}$ 与 λ 进行比较，作出判断：当 $\|\hat{K}\|>\lambda$(或 $\hat{K}>\lambda$) 时拒绝 H_0，否则认为接受 H_0。	
两类错误	第一类错误	当 H_0 为真时，而样本值却落入了拒绝域，应当否定 H_0。这时，我们把客观上 H_0 成立判为 H_0 为不成立（即否定了真实的假设），称这种错误为“弃真错误”或第一类错误，记 α 为犯此类错误的概率，即：P{拒绝 $H_0\|H_0$ 为真}= α；
	第二类错误	当 H_1 为真时，而样本值却落入了接受域，按照我们规定的检验法则，应当接受 H_0。这时，我们把客观上 H_0 不成立判为 H_0 成立（即接受了不真实的假设），称这种错误为“取伪错误”或第二类错误，记 β 为犯此类错误的概率，即： P{接受 $H_0\|H_1$ 为真}= β。
	两类错误的关系	人们当然希望犯两类错误的概率同时都很小。但是，当容量 n 一定时，α 变小，则 β 变大；相反地，β 变小，则 α 变大。取定 α 要想使 β 变小，则必须增加样本容量。

2. 单正态总体均值和方差的假设检验

条件	原假设	检验统计量	统计量分布	拒绝域
已知 σ^2	$H_0:\mu=\mu_0$	$Z=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$N(0,1)$	$\|z\|>z_{\alpha/2}$
	$H_0:\mu\le\mu_0$			$z>z_\alpha$

	$H_0:\mu\ge\mu_0$			$z<-z_\alpha$
未知σ^2	$H_0:\mu=\mu_0$	$T=\frac{\bar{x}-\mu_0}{S/\sqrt{n}}$	$t(n-1)$	$\lvert t\rvert>t_{\alpha/2}(n-1)$
	$H_0:\mu\le\mu_0$			$t>t_\alpha(n-1)$
	$H_0:\mu\ge\mu_0$			$t<-t_\alpha(n-1)$
未知μ	$H_0:\sigma^2=\sigma^2$	$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}$	$\chi^2(n-1)$	$\chi^2<\chi^2_{1-\alpha/2}(n-1)$或 $\chi^2>\chi^2_{\alpha/2}(n-1)$
	$H_0:\sigma^2\le\sigma_0^2$			$\chi^2>\chi^2_\alpha(n-1)$
	$H_0:\sigma^2\ge\sigma_0^2$			$\chi^2<\chi^2_{1-\alpha}(n-1)$
已知μ （少见）	$H_0:\sigma^2=\sigma^2$	$\chi^2=\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}$	$\chi^2(n)$	$\chi^2<\chi^2_{1-\alpha/2}(n)$或 $\chi^2>\chi^2_{\alpha/2}(n)$
	$H_0:\sigma^2\le\sigma_0^2$			$\chi^2>\chi^2_\alpha(n)$
	$H_0:\sigma^2\ge\sigma_0^2$			$\chi^2<\chi^2_{1-\alpha}(n)$