# 概率论与数理统计公式集锦

## 一、随机事件与概率

| 公式名称 | 公式表达式 |
|---|---|
| 德摩根公式 | $\overline{A\cup B}=\overline{A}\cap\overline{B}$，$\overline{A\cap B}=\overline{A}\cup\overline{B}$ |
| 古典概型 | $P(A)=\frac{m}{n}=\frac{A包含的基本事件数}{基本事件总数}$ |
| 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)<br>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) |
| 条件概率公式与乘法公式 | $P(B\|A)=\frac{P(AB)}{P(A)}$ $P(AB)=P(A)P(B\|A)=P(B)P(A\|B)$<br>$P(ABC)=P(A)P(B\|A)P(C\|AB)$ |
| 全概率公式 | $P(A)=\sum_{i=1}^{n}P(B_i)P(A\|B_i)$ |
| 贝叶斯公式（逆概率公式） | $P(B_i\|A)=\frac{P(B_i)P(A\|B_i)}{\sum_{i=1}^{n}P(B_i)P(A\|B_i)}$ |
| 两个事件相互独立 | $P(AB)=P(A)P(B)$；$P(B\|A)=P(B)$；$P(B\|A)=P(B\|\overline{A})$； |

## 二、随机变量及其分布

1、分布函数

$$F(x)=P(X\le x)=\begin{cases}\sum_{x_k\le x}P(X=x_k)\\ \int_{-\infty}^{x}f(t)dt\end{cases},\quad P(a<X\le b)=F(b)-F(a)$$

2、离散型随机变量及其分布

| 分布名称 | 分布律 |
|---|---|
| 0–1 分布 $X\sim b(1,p)$ | $P(X=k)=p^k(1-p)^{1-k},\quad k=0,1$ |
| 二项分布 $X\sim b(n,p)$ | $P(X=k)=C_n^k p^k(1-p)^{n-k},\quad k=0,1,\cdots,n$ |
| 泊松分布 $X\sim P(\lambda)$ | $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,\cdots$ |

3、续型随机变量及其分布

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

4、随机变量函数 Y=g(X)的分布

离散型：$P(Y=y_i)=\sum_{g(x_j)=y_i}p_j, i=1,2,\cdots$，

连续型：①分布函数法，②公式法 $f_Y(y)=f_X(h(y))\cdot|h'(y)|$$(x=h(y)$单调)

## 三、多维随机变量及其分布

1、离散型二维随机变量及其分布

分布律：$P(X=x_i,Y=y_j)=p_{ij}, i,j=1,2,\cdots$ 分布函数 $F(X,Y)=\sum_{x_i\le x}\sum_{y_i\le y}p_{ij}$

边缘分布律：$p_{i\cdot}=P(X=x_i)=\sum_j p_{ij}$　$p_{\cdot j}=P(Y=y_j)=\sum_i p_{ij}$

条件分布律：$P(X=x_i|Y=y_j)=\frac{p_{ij}}{p_{\cdot j}}, i=1,2,\cdots$，$P(Y=y_j|X=x_i)=\frac{p_{ij}}{p_{i\cdot}}, j=1,2,\cdots$

2、连续型二维随机变量及其分布

①分布函数及性质

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

性质：$F(+\infty,+\infty)=1,\frac{\partial^2F(x,y)}{\partial x\partial y}=f(x,y),\ P((x,y)\in G)=\iint_G f(x,y)dxdy$

②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：$f_X(x)=\int_{-\infty}^{+\infty}f(x,v)dv$

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

③条件概率密度

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

3、随机变量的独立性

随机变量 X、Y 相互独立$\Leftrightarrow F(x,y)=F_X(x)F_Y(y)$，

离散型：$p_{ij}=p_{i.}p_{.j}$，连续型：$f(x,y)=f_X(x)f_Y(y)$

4、二维随机变量和函数的分布

离散型：$P(Z=z_k)=\sum_{x_i+y_j=z_k}P(X=x_i,Y=y_j)$

连续型：$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)dx=\int_{-\infty}^{+\infty}f(z-y,y)dy$

## 四、随机变量的数字特征

1、数学期望

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

②性质：$E(C)=C,\ E[E(X)]=E(X)$，$E(CX)=CE(X)$，$E(X\pm Y)=E(X)\pm E(Y)$

$E(aX\pm b)=aE(X)\pm b$，当 X、Y 相互独立时：$E(XY)=E(X)E(Y)$

2、方差

①定义：$D(X)=E[(X-E(X))^2]=E(X^2)-E^2(X)$

②性质：$D(C)=0$，$D(aX\pm b)=a^2D(X)$，$D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$

当 X、Y 相互独立时：$D(X\pm Y)=D(X)+D(Y)$

3、协方差与相关系数

①协方差：$Cov(X,Y)=E(XY)-E(X)E(Y)$，当 X、Y 相互独立时：$Cov(X,Y)=0$

②相关系数：$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$，当 X、Y 相互独立时：$\rho_{XY}=0$ (X, Y 不相关)

③协方差和相关系数的性质：$Cov(X,X)=D(X)$，$Cov(X,Y)=Cov(Y,X)$

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

4、常见随机变量分布的数学期望和方差

| 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 $P(\lambda)$ | $\lambda$ | $\lambda$ |
| 均匀分布 $U(a,b)$ | $\frac{a+b}{2}$ | $\frac{(b-a)^2}{12}$ |
| 正态分布 $N(\mu,\sigma^2)$ | $\mu$ | $\sigma^2$ |
| 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 $P\{|X-E(X)|\ge\varepsilon\}\le\frac{D(X)}{\varepsilon^2}$

2、大数定律：①切比雪夫大数定律：若 $X_1\cdots X_n$ 相互独立，

$E(X_i)=\mu_i,D(X_i)=\sigma_i^2$ 且 $\sigma_i^2\le C$，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伯努利大数定律：设 $n_A$ 是 $n$ 次独立试验中事件 A 发生的次数，$p$ 是事件 A 在每次试验中发生的概率，则$\forall\varepsilon>0$，有：$\lim_{n\to\infty}P\left(\left|\frac{n_A}{n}-p\right|<\varepsilon\right)=1$

③辛钦大数定律：若 $X_1,\cdots,X_n$ 独立同分布，且 $E(X_i)=\mu$，则 $\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow[n\to\infty]{P}\mu$

3、中心极限定理

①列维—林德伯格中心极限定理：独立同分布的随机变量 $X_i(i=1,2,\cdots)$，均值为 $\mu$，方差为 $\sigma^2>0$，当 n 充分大时有：$Y_n=(\sum_{k=1}^{n}X_k-n\mu)\Big/\sqrt{n}\sigma\xrightarrow{\sim}N(0,1)$

②棣莫弗—拉普拉斯中心极限定理：随机变量 $X\sim B(n,p)$，则对任意 $x$ 有：

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

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

# 六、数理统计的基本概念

1、总体和样本的分布函数

设总体 $X$ □ $F(x)$，则样本的联合分布函数 $F(x_1,x_2\cdots x_n)=\prod_{k=1}^{n}F(x_k)$

2、统计量

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

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

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

3、三大抽样分布

(1) $\chi^2$ 分布：设随机变量 $X_i$ □ $N(0,1)$ $(i=1,2,\cdots,n)$ 且相互独立，则称统计量 $\chi^2=X_1^2+X_2^2+\cdots X_n^2$ 服从自由度为 $n$ 的 $\chi^2$ 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，则称统计量：$T=\frac{X}{\sqrt{Y/n}}$ 服从自由度为 $n$ 的 $t$ 分布，记为 $T\sim t(n)$

性质：① $E(T)=0\ (n>1),D(T)=\frac{n}{n-2}\ (n>2)$ ② $\lim_{n\to\infty}f_n(x)=\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

(3) $F$ 分布：设随机变量 $X\sim\chi^2(m),Y\sim\chi^2(n)$，且 $X$ 与 $Y$ 独立，则称统计量 $F(m,n)=\frac{X/m}{Y/n}$ 服从第一自由度为 $m$，第二自由度为 $n$ 的 $F$ 分布，记为 $F\sim F(m,n)$，性质：设 $F\sim F(m,n)$，则 $\frac{1}{F}\sim F(n,m)$

# 七、参数估计

1. 参数估计

①定义：用 $\hat{\theta}(X_1,X_2,\mathrm{L}\ ,X_n)$ 估计总体参数 $\theta$，称 $\hat{\theta}(X_1,X_2,\mathrm{L}\ ,X_n)$ 为 $\theta$ 的估计量，相应的 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 为总体 $\theta$ 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：

基本思想：用样本矩来估计相应的总体矩

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 $x_1,x_2,\mathrm{L}\ ,x_n$ 为总体 X 的 $n$ 个样本值，用样本矩代替 $\mu_i$，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 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$。

**注意：分布中有几个未知参数，就求到几阶矩。**

3. 点估计中的极大似然估计

设 $X_1,X_2,\mathrm{L}\ X_n$ 取自 $X$ 的样本，设 $X\sim f(x,\theta)$ 或 $X\sim P(x,\theta)$，求法步骤：

①似然函数：$L(\theta)=\prod_{i=1}^{n}f(x_i,\theta)(\text{连续型})\text{或}L(\theta)=\prod_{i=1}^{n}P_i(x_i,\theta)(\text{离散型})$

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

③解方程：$\frac{\partial\ln L}{\partial\theta_1}=0,\mathrm{L}\ ,\frac{\partial\ln L}{\partial\theta_k}=0$，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 $\hat{\theta}=\hat{\theta}(x_1,x_2,\mathrm{L}\ ,x_n)$ 为未知参数 $\theta$ 的估计量。若 E（$\hat{\theta}$）=$\theta$，则称 $\hat{\theta}$ 为 $\theta$ 的无偏估计量。 |
| | 有效性 | 设 $\hat{\theta}_1=\hat{\theta}_1(x_1,x_2,\mathrm{L}\ ,x_n)$ 和 $\hat{\theta}_2=\hat{\theta}_2(x_1,x_2,\mathrm{L}\ ,x_n)$ 是未知参数 $\theta$ 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 $\theta$ 的一串估计量，如 $\forall\varepsilon>0$，有 $\lim_{n\to\infty}P(\lvert\hat{\theta}_n-\theta\rvert>\varepsilon)=0$ 则称 $\hat{\theta}_n$ 为 $\theta$ 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 $\sigma^2$ | $\mu$ | $Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ | $N(0,1)$ | $\left(\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$ |
| 未知 $\sigma^2$ | $\mu$ | $T=\frac{\overline{X}-\mu}{S/\sqrt{n}}$ | $t(n-1)$ | $\left(\bar{x}-t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}},\bar{x}+t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)$ |
| 已知 $\mu$ | $\sigma^2$ | $\chi^2=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$ | $\chi^2(n)$ | $\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)},\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right)$ |
| 未知 $\mu$ | $\sigma^2$ | $\chi^2=\frac{(n-1)S^2}{\sigma^2}$ | $\chi^2(n-1)$ | $\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$ |

# 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。<br>小概率事件的概率就是显著性水平 α，常取 α=0.05，0.01 或 0.10。 |
| 基本步骤 | | ①提出原假设 $H_0$；②选择检验统计量 $g(X_1,\mathrm{L}\ ,X_n)$；③对于 α 查表找分位数 λ，使 $P(g(X_1,\mathrm{L}\ ,X_n)\in W)=\alpha$，从而定出拒绝域 $W$；④由样本观测值计算统计量实测值 $g(x_1,\cdots,x_n)$；并作出判断：当实测值落入 $W$ 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 $\alpha$ 为犯此类错误的概率，即：P{拒绝 $H_0$\|$H_0$ 为真}=$\alpha$； |
| | 第二类错误 | 当 $H_1$ 为真时，而样本值却落入了接受域，应接受 $H_0$。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 $\beta$ 为犯此类错误的概率，即：P{接受 $H_0$\|$H_1$ 为真}=$\beta$。 |
| | 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，$\alpha$ 变小，则 $\beta$ 变大；相反地，$\beta$ 变小，则 $\alpha$ 变大。取定 $\alpha$ 要想使 $\beta$ 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 $\sigma^2$ | $H_0:\mu=\mu_0$ | $Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ | $N(0,1)$ | $\lvert z\rvert>z_{\alpha/2}$ |
| | $H_0:\mu\le\mu_0$ | | | $z>z_\alpha$ |
| | $H_0:\mu\ge\mu_0$ | | | $z<-z_\alpha$ |
| 未知 $\sigma^2$ | $H_0:\mu=\mu_0$ | $T=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$ | $t(n-1)$ | $\lvert t\rvert>t_{\alpha/2}(n-1)$ |
| | $H_0:\mu\le\mu_0$ | | | $t>t_\alpha(n-1)$ |
| | $H_0:\mu\ge\mu_0$ | | | $t<-t_\alpha(n-1)$ |
| 未知 $\mu$ | $H_0:\sigma^2=\sigma^2$ | $\chi^2=\frac{(n-1)S^2}{\sigma_0^2}$ | $\chi^2(n-1)$ | $\chi^2<\chi^2_{1-\alpha/2}(n-1)$ 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 $\mu$（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 $\chi^2>\chi^2_{\alpha/2}(n)$ |
| | $H_0:\sigma^2\le\sigma_0^2$ | | | $\chi^2>\chi^2_\alpha(n)$ |
| | $H_0:\sigma^2\ge\sigma_0^2$ | | | $\chi^2<\chi^2_{1-\alpha}(n)$ |