考研数学公式汇总（概率统计篇）

1.概率基本公式

(1)逆事件的概率 $P(\overline{A})=1-P(A)$.

注： 正面直接求概率困难时可考虑此公式, 比如涉及"至少、至多"等字眼.

(2) 加法公式 $P(A\cup B)=P(A)+P(B)-P(AB)$;

$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$.

注： 超过3个事件的加法公式往往会有两两互斥的条件.

(3) 减法公式 $P(A-B)=P(A)-P(AB)=P(A\overline{B})$.

注： 考减法公式是考试的重点.

(4) 条件概率 若$P(A)>0$, 称在A发生的条件下, B发生的概率为条件概率,

记为$P(B|A)$, 且$P(B|A)=\dfrac{P(AB)}{P(A)}$.

注： 条件概率$P(B|A)$也是概率, 满足概率的一切性质与公式, 如

$P(\overline{B}|A)=1-P(B|A)$; $P(B-C|A)=P(B|A)-P(BC|A)=P(B\overline{C}|A)$.

(5) 乘法公式 如果$P(A)>0$, 则$P(AB)=P(A)\cdot P(B|A)$.

(6) 全概率公式 若$A_1\cup A_2\cup\cdots\cup A_n=\Omega$, 且$A_i\cap A_j=\Phi, 1\le i\ne j\le n$,

则对任一事件B, 有$P(B)=\sum\limits_{i=1}^{n}P(A_i)P(B|A_i)$.

注： 如果某个事件B的发生总是与某些原因或前一阶段的某些结果A_i有关,
则总是使用全概率公式把各种导致B发生的可能性（概率）加起来求$P(B)$.

(7) 贝叶斯公式 若$A_1\cup A_2\cup\cdots\cup A_n=\Omega$, 且$A_i\cap A_j=\Phi, 1\le i\ne j\le n$,

则对任一事件B, 只要$P(B)>0$, 则$P(A_j|B)=\dfrac{P(A_j)P(B|A_j)}{\sum\limits_{i=1}^{n}P(A_i)P(B|A_i)}$.

注： 如果已知B发生了, 去探求是某原因A_j导致发生的可能性（概率）$P(A_j|B)$,
则总是使用贝叶斯公式看这一原因占总的原因的比例.

2. 独立与互斥、包含的关系

设$0<P(A)<1, 0<P(B)<1$, 如果A与B互斥或存在包含关系, 则A与B不独立.

3.常见的分布

1. $0-1$分布

如果随机变量X的分布律为$P\{X=k\}=p^k(1-p)^{1-k},k=0,1.$

则称X服从参数为$p(0<p<1)$的$0-1$分布，记为$X\sim B(1,p)$.

2. 二项分布

如果随机变量X的分布律为$P\{X=k\}=C_n^k p^k(1-p)^{n-k},k=0,1,\cdots,n.$

则称X服从参数为$n,p(0<p<1)$的二项分布，记为$X\sim B(n,p)$.

注：（1）n次伯努利试验中试验成功的次数X服从二项分布；

（2）对$X\sim B(n,p)$，最可能发生（成功）的次数k满足$(n+1)p-1\le k\le(n+1)p$.

3. 泊松分布

如果随机变量X的分布律为$P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!},k=0,1,2,\cdots$

则称X服从参数为$\lambda(\lambda>0)$的泊松分布，记为$X\sim P(\lambda)$.

4. 几何分布

如果随机变量X的分布律为$P\{X=k\}=(1-p)^{k-1}p,k=1,2,\cdots$

则称X服从参数为$p(0<p<1)$的几何分布，记为$X\sim G(p)$.

注：伯努利试验中首次成功所需的试验次数X服从几何分布.

5.均匀分布

如果随机变量X的概率密度为$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & 其他\end{cases}$,

则称X服从(a,b)上的均匀分布，记为$X\sim U(a,b)$.X的分布函数为$F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\le x<b\\ 1, & x\ge b\end{cases}$.

注：若$X\sim U(a,b)$，对$a\le c<d\le b$，则$P\{c<X<d\}=\dfrac{d-c}{b-a}$.

6. 指数分布

如果随机变量X的概率密度为$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & 其他\end{cases}$，其中$\lambda>0$为参数；

则称X服从参数为λ的指数分布，记为$X\sim E(\lambda)$.X的分布函数为$F(x)=\begin{cases}1-e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}$.

注：若$X\sim E(\lambda)$，则

1° 对$\forall a>0$，则$P\{X\ge a\}=e^{-\lambda a}$；

2° 对$\forall t,s>0$，则$P\{X\ge t+s|X\ge s\}=P\{X\ge t\}$.

7. 正态分布

如果随机变量X的概率密度为:$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty.$

则称X服从参数为μ,σ^2的正态分布,记为$X\sim N(\mu,\sigma^2)$.

特别地,当$\mu=0,\sigma=1$时称为**标准正态分布,** 记为$X\sim N(0,1)$;

概率密度$\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},-\infty<x<+\infty$;分布函数$\Phi(x)=\int_{-\infty}^{x}\varphi(t)dt=\int_{-\infty}^{x}\dfrac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt.$

注: 1°若$X\sim N(\mu,\sigma^2)$,则$\dfrac{X-\mu}{\sigma}\sim N(0,1)$;(标准化)

$2^{\circ}\ \Phi(-x)=1-\Phi(x),\Phi(0)=\dfrac{1}{2}$;

3°若$X\sim N(\mu,\sigma^2)$,则$aX+b\sim N(a\mu+b,a^2\sigma^2)$;

4°若X,Y分别服从正态分布,且相互独立,则$aX+bY$服从正态分布.

4. 两个常见的二维连续型随机变量

1.二维均匀

(X,Y)在平面区域D上服从均匀分布,则

$$f(x,y)=\begin{cases}\dfrac{1}{S_D},(X,Y)\in D\\ 0,\qquad 其他\end{cases},\ 其中S_D是D的面积.$$

注: 设(X,Y)在区域D上服从均匀分布,若$G\subset D$,则$P\{(X,Y)\in G\}=\dfrac{S_G}{S_D}$;

2.二维正态

$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2;\rho)$.其中$\mu_1=EX,\mu_2=EY,\sigma_1^2=DX,\sigma_2^2=DY;\rho\in(-1,1)$.

注: (1) $X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$,反之不对（独立时可以）;

(2) X,Y的条件分布都是正态分布;

(3) $aX+bY$服从正态分布;

(4) X,Y独立$\Leftrightarrow X,Y$不相关,即$\rho=0$.

5.期望

(1) 一维离散型

设离散型随机变量X的分布律为$P\left\{X=x_i\right\}=p_i\left(i=1,2,\cdots\right)$,

$Y=g\left(X\right)$是X的函数, 则$EX=\sum_{i=1}^{\infty}x_ip_i$; $Eg\left(X\right)=\sum_{i=1}^{\infty}g\left(x_i\right)p_i$.

(2) 一维连续型

设连续型随机变量X的概率密度为$f\left(x\right)$,

$Y=g\left(X\right)$是X的函数, 则$EX=\int_{-\infty}^{+\infty}xf\left(x\right)dx$; $Eg\left(X\right)=\int_{-\infty}^{+\infty}g\left(x\right)f\left(x\right)dx$.

(3) 二维离散型

设二维离散型随机变量$\left(X,Y\right)$的联合分布为$P\left\{X=x_i,Y=y_j\right\}=p_{ij}$

$\left(i,j=1,2,\cdots\right)$, $Z=g\left(X,Y\right)$是$\left(X,Y\right)$的函数, 则$Eg\left(X,Y\right)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}g\left(x_i,y_j\right)p_{ij}$.

(4) 二维连续型

设二维连续型随机变量$\left(X,Y\right)$的联合概率密度为$f\left(x,y\right)$,

$Z=g\left(X,Y\right)$是$\left(X,Y\right)$的函数, 则$Eg\left(X,Y\right)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g\left(x,y\right)f\left(x,y\right)dxdy$.

(5) 性质

$1^{\circ}\ Ec=c$;

$2^{\circ}\ E\left(aX+c\right)=aEX+c$;

$3^{\circ}\ E\left(X\pm Y\right)=EX\pm EY$;

4°若X,Y独立, 则$E\left(XY\right)=EX\cdot EY$.

6.方差

(1) 定义

$DX=E\left(X-EX\right)^2=EX^2-\left(EX\right)^2$.

(2) 性质

$1^{\circ}\ DX\geq 0, EX^2=\left(EX\right)^2+DX$;

$2^{\circ}\ Dc=0$;

$3^{\circ}\ D\left(aX+b\right)=a^2DX$;

$4^{\circ}\ D\left(X\pm Y\right)=DX+DY\pm 2Cov\left(X,Y\right)$;

5°若X,Y独立, 则$D\left(X\pm Y\right)=DX+DY$,

$D\left(XY\right)=DXDY+DX\left(EY\right)^2+DY\left(EX\right)^2$.

7.常用分布的数学期望和方差

1° 如果 $X \sim B(1,p)$, 则 $EX=p, DX=p(1-p)$;

2° 如果 $X \sim B(n,p)$, 则 $EX=np, DX=np(1-p)$;

3° 如果 $X \sim P(\lambda)$, 则 $EX=\lambda, DX=\lambda$;

4° 如果 $X \sim G(p)$, 则 $EX=\dfrac{1}{p}, DX=\dfrac{1-p}{p^2}$;

5° 如果 $X \sim U(a,b)$, 则 $EX=\dfrac{a+b}{2}, DX=\dfrac{(b-a)^2}{12}$;

6° 如果 $X \sim E(\lambda)$, 则 $EX=\dfrac{1}{\lambda}, DX=\dfrac{1}{\lambda^2}$;

7° 如果 $X \sim N(\mu,\sigma^2)$, 则 $EX=\mu, DX=\sigma^2$;

8° 如果 $X \sim N(0,1)$, 则 $E|X|=\sqrt{\dfrac{2}{\pi}}, D|X|=1-\dfrac{2}{\pi}$.

8.协方差

(1) 定义

$Cov(X,Y)=E[(X-EX)(Y-EY)]=E(XY)-EX\cdot EY.$

(2) 性质

$1^\circ\ Cov(X,Y)=Cov(Y,X), Cov(X,X)=DX$;

$2^\circ\ Cov(X,c)=0$;

$3^\circ\ Cov(aX,bY)=abCov(X,Y)$;

$4^\circ\ Cov(aX_1+bX_2,cY_1+dY_2)=$

$acCov(X_1,Y_1)+adCov(X_1,Y_2)+bcCov(X_2,Y_1)+bdCov(X_2,Y_2).$

9.相关系数

(1) 定义

$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{DX}\cdot\sqrt{DY}}$, 如果 $\rho_{XY}=0$, 称 X 和 Y 不相关.

(2) 性质

$1^\circ\ \rho_{XY}=\rho_{YX}$; $\rho_{XX}=1$;

$2^\circ\ |\rho_{XY}|\le 1$;

$3^\circ\ |\rho_{XY}|=1 \Leftrightarrow$ 存在 a,b 使 $P\{Y=aX+b\}=1$;

4° 如果 $Y=aX+b$, 则 $\rho_{XY}=\begin{cases}1, a>0\\-1, a<0\end{cases}$.

10.大数定律

1.依概率收敛

对随机变量序列$X_1, X_2, \cdots, X_n, \cdots$,和常数$a$,如果对任意的$\varepsilon > 0$,有$\lim\limits_{n\to\infty} P\{|X_n - a| < \varepsilon\} = 1$,

则称随机变量序列$X_1, X_2, \cdots, X_n, \cdots$,依概率收敛于$a$,记为$X_n \xrightarrow{P} a$.

2.切比雪夫大数定律

设$X_1, X_2, \cdots, X_n, \cdots$,独立,期望$EX_k$,方差$DX_k$都存在,方差$DX_k$有一致上界,$k = 1, 2, \cdots$,

则对任意的$\varepsilon > 0$,有$\lim\limits_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} EX_i\right| < \varepsilon\right\} = 1.$

3.伯努利大数定律

设X是n重伯努利试验中事件A发生的次数,每次试验事件A发生的概率为p,

即$X \sim B(n, p)$,则对任意的$\varepsilon > 0$,有$\lim\limits_{n\to\infty} P\left\{\left|\frac{X}{n} - p\right| < \varepsilon\right\} = 1.$

4.辛钦大数定律

设$X_1, X_2, \cdots, X_n, \cdots$,独立同分布,期望$EX_k = \mu$存在,则对任意的$\varepsilon > 0$,有$\lim\limits_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1.$

11.中心极限定理

1.列维—林德伯格中心极限定理

设$X_1, X_2, \cdots, X_n, \cdots$,独立同分布,期望$EX_k = \mu$,方差$DX_k = \sigma^2$都存在,

则对任意的x,有$\lim\limits_{n\to\infty} P\left\{\frac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).$

2.拉普拉斯中心极限定理

设$X \sim B(n, p)$,则对任意的x,有$\lim\limits_{n\to\infty} P\left\{\frac{X - np}{\sqrt{np(1-p)}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).$

12.三大抽样分布

1.$\chi^2(n)$分布

(1) 定义： 设$X_1, X_2, \cdots, X_n$相互独立且都服从标准正态$N(0,1)$,

则$X_1^2 + X_2^2 + \cdots + X_n^2$服从自由度为$n$的$\chi^2$分布, 记为$X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$.

(2) 上α分位点

对于给定的$\alpha(0<\alpha<1)$, 称满足$P\{\chi^2(n) > \chi^2_\alpha(n)\} = \int_{\chi^2_\alpha(n)}^{+\infty} f(x)dx = \alpha$

（$f(x)$是$\chi^2(n)$的概率密度）的数$\chi^2_\alpha(n)$为$\chi^2(n)$的上α分位点.

(3) χ^2分布的性质

若$X \sim \chi^2(n)$, 则$EX = n, DX = 2n$;

若$X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$, 且$X, Y$独立, 则$X + Y \sim \chi^2(n_1 + n_2)$.

2.t分布

(1) 定义： 设$X \sim N(0,1), Y \sim \chi^2(n)$, 且$X, Y$独立,

则$\dfrac{X}{\sqrt{Y/n}}$服从自由度为n的t分布, 记为$\dfrac{X}{\sqrt{Y/n}} \sim t(n)$.

(2) 上α分位点

对于给定的$\alpha(0<\alpha<1)$, 称满足$P\{t(n) > t_\alpha(n)\} = \int_{t_\alpha(n)}^{+\infty} f(x)dx = \alpha$

（$f(x)$是$t(n)$的概率密度）的数$t_\alpha(n)$为$t(n)$的上α分位点.

(3) t分布的性质

t分布的概率密度$f(x)$是偶函数, 故$t_{1-\alpha}(n) = -t_\alpha(n)$, 且当自由度$n$充分大时, $t(n)$分布近似于$N(0,1)$;

$t \sim t(n)$, 则$t^2 \sim F(1,n)$.

3.F分布

(1) 定义： 设$X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$, 且$X, Y$独立,

则$\dfrac{X/n_1}{Y/n_2}$服从第一自由度为n_1, 第二自由度为n_2的F分布, 记为$\dfrac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$.

(2) 上α分位点

对于给定的$\alpha(0<\alpha<1)$, 称满足$P\{F(n_1,n_2) > F_\alpha(n_1,n_2)\} = \int_{F_\alpha(n_1,n_2)}^{+\infty} f(x)dx = \alpha$

（$f(x)$是$F(n_1,n_2)$的概率密度）的数$F_\alpha(n_1,n_2)$为$F(n_1,n_2)$的上α分位点.

(3) F分布的性质

若$F \sim F(n_1,n_2)$, 则$\dfrac{1}{F} \sim F(n_2,n_1)$;

若$F \sim F(n_1,n_2)$, 则$F_{1-\alpha}(n_1,n_2) = \dfrac{1}{F_\alpha(n_2,n_1)}$.

13.矩估计的求法

原理：用样本矩替换总体矩——$\alpha_k = A_k$,即$\frac{1}{n}\sum_{i=1}^{n} X_i^k = EX^k$.

步骤：对一个未知参数的情形　令$\overline{X} = EX$.

对两个未知参数的情形　令$\begin{cases} \overline{X} = EX \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 = EX^2 \end{cases}$或$\begin{cases} \overline{X} = EX \\ \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2 = DX \end{cases}$.

14.最大似然估计的求法

步骤：a.写出样本的似然函数

$L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{i=n} f(x_i, \theta)$[**连续型**]

$L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{i=n} p(x_i, \theta)$[**离散型**]

b.取对数得$\ln L$

c.求导$\frac{d \ln L}{d\theta} = 0$,解出$\theta$即可.

若$\frac{d \ln L}{d\theta} = 0$无解,即$\ln L$单调,则应该用定义法找出$\theta$的最大似然估计量.

15.估计量的评价标准

（1）无偏性

若$E\hat{\theta} = \theta$,则称$\hat{\theta}$是θ的无偏估计量.

（2）有效性

设$\hat{\theta}_1, \hat{\theta}_2$都是$\theta$的无偏估计量,若$D\hat{\theta}_1 < D\hat{\theta}_2$,则称$\hat{\theta}_1$比$\hat{\theta}_2$更有效.

（3）一致性

若对任意的$\varepsilon > 0$,有$\lim_{n\to\infty} P\left\{\left|\hat{\theta} - \theta\right| < \varepsilon\right\} = 1$,即$\hat{\theta} \xrightarrow{P} \theta$,则称$\hat{\theta}$是$\theta$的一致估计量.

16. 求置信区间的步骤

（1）构造统计量T并确定其分布;

（2）给定α,确定常数a, b,使得$P\{a < T < b\} = 1 - \alpha$;

（3）由（2）$a < T < b$反解出θ的范围:$\hat{\theta}_1 < \theta < \hat{\theta}_2$,得置信区间$\left(\hat{\theta}_1, \hat{\theta}_2\right)$.