概率论公式整理

程铁平 ypcheng@bjtu.edu.cn

收录原则：1）容易记的不收录（如集合运算法则，各种线性双线性公式）2）不够常用的不收录

第一章

加法公式：对任意两个事件 A, B, 有

$$P(A\cup B)=P(A)+P(B)-P(AB)$$

条件概率

$$P(B\mid A)=\frac{P(AB)}{P(A)}$$

乘法公式

$$P(AB)=P(A)P(B\mid A)$$

全概率公式

$$P(A)=\sum_{i=1}^{n}P(AB_i)=\sum_{i=1}^{n}P(B_i)\cdot P(A\mid B_i)$$

Bayes 公式

$$P(B_k\mid A)=\frac{P(AB_k)}{P(A)}=\frac{P(B_k)P(A\mid B_k)}{\sum_{i=1}^{n}P(B_i)P(A\mid B_i)}$$

第二章

最常用离散型分布

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

Poisson 分布 $\pi(\lambda)$	$P(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,\cdots$	λ	λ

最常用连续型分布

分布名称	概率密度函数	分布函数	期望	方差
均匀分布 $U(a,b)$	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases}$	$F(x)=\begin{cases}0, & x\le a\\ \dfrac{x-a}{b-a}, & a<x<b\\ 1, & x\ge b\end{cases}$		$\dfrac{(b-a)^2}{12}$
正态分布 $N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$F(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t$		
指数分布 -λ版本	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & \text{其他}\end{cases}$	$F(x)=\begin{cases}1-e^{-\lambda x}, & x>0\\ 0, & \text{其他}\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
指数分布 -θ版本	$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0\\ 0, & \text{其他}\end{cases}$	$F(x)=\begin{cases}1-e^{-\frac{x}{\theta}}, & x>0\\ 0, & \text{其他}\end{cases}$	θ	θ^2

* $N(0,1)$ — 标准正态分布

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\mathrm{d}t$$

两个常用积分

$$\int_{-\infty}^{\infty}e^{-ax^2}\mathrm{d}x=\sqrt{\frac{\pi}{a}}$$

$$\int_0^{\infty} x^n e^{-x} \mathrm{d}x = n!$$

第三章

二维连续随机变量根据联合密度函数求概率

$P\{(x,y)\in G\} = \iint\limits_G f(x,y)dxdy$。本式为二重积分，实际计算时请化成累次积分（下同）。

二维连续随机变量根据联合密度函数求联合分布函数

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

二维连续随机变量根据联合分布函数求联合密度函数

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

二维连续随机变量根据联合密度函数求边缘密度函数

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx$$

二维连续随机变量根据联合密度函数求条件密度函数

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

二维连续随机变量根据一个边缘密度函数和条件密度函数，求联合密度函数

同样基于上面的公式

二维随机变量根据联合分布函数求边缘分布函数

$$F_X(x) = F(x,\infty)$$

$$F_Y(y) = F(\infty,y)$$

Z=X+Y 根据 X,Y 联合**密度**函数求 Z **密度**函数

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x,z-x)dx$$

Z=X/Y 根据 X,Y 联合**密度**函数求 Z **密度**函数

$$f_{X/Y}(z)=\int_{-\infty}^{\infty}|y|f(yz,y)dy$$

Z=Y/X 根据 X,Y 联合密度函数求 Z 密度函数

$$f_{Y/X}(z)=\int_{-\infty}^{\infty}|x|f(x,xz)dx$$

Z=max(X,Y)且 X,Y 独立，根据各自分布函数求 Z 分布函数

$$F_Z(z)=F_X(z)F_Y(z)$$

第四章

离散型随机变量的数学期望

$E(X)=\sum_{k=1}^{n}x_k p_k$，其中 $x_1,\cdots,x_n$ 为 X 的所有可能值，$p_1,\cdots,p_n$ 为其相应的概率。其中 n 可以是无穷大。

连续型随机变量的数学期望

$E(X)=\int_{-\infty}^{+\infty}xf(x)dx$，其中 $f(x)$ 为密度函数

二维连续随机变量根据联合密度函数求各种常用期望值

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy$$

$$E(X^2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^2 f(x,y)dxdy$$

$$E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dxdy$$

$$E(Y^2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y^2 f(x,y)dxdy$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)dxdy$$

求方差

$$D(X)=E(X^2)-E^2(X)$$

求协方差

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)$$

求相关系数

$$\rho_{X,Y}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

方差与协方差关系

$$D(X)=\mathrm{Cov}(X,X)$$

$$D(aX+bY)=a^2D(X)+b^2D(Y)+2ab\mathrm{Cov}(X,Y)$$

车比雪夫定理

$P(|X-\mu|\geq k\sigma)\leq\dfrac{1}{k^2}$，其中 μ,σ 分别为期望和标准差。

第六章

样本方差，它是总体方差的无偏估计

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

样本 k 阶原点矩

$$\boldsymbol{A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k\quad(k=1,2,\ldots)}$$

样本 k 阶中心矩

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

第四章“求方差公式”的样本矩版本

$$B_2=A_2-A_1^2$$

$\chi^2(n)$ 的方差为 2n

设 $Z_1,Z_2,\cdots$ 服从 $N(0,1)$ 且相互独立，令 $\overline{Z}=\dfrac{Z_1+\cdots+Z_n}{n}$。则

$Z_1^2+\cdots+Z_n^2$ 服从 $\chi^2(n)$

$\dfrac{Z_1}{\sqrt{(Z_2^2+\cdots+Z_{n+1}^2)/n}}$ 服从 $t(n)$

$\dfrac{(Z_1^2+\cdots+Z_m^2)/m}{(Z_{m+1}^2+\cdots+Z_{m+n}^2)/n}$ 服从 $F(m,n)$

$\sqrt{n}\overline{Z}\sim N(0,1)$

$$\sum_{i=1}^{n}(Z_i-\overline{Z})^2 : \ \chi^2(n-1)$$

设 X_1, X_2, L 服从 $N(\mu,\sigma^2)$且相互独立，令 $\overline{X}=\dfrac{X_1+\mathrm{L}\ +X_n}{n}$。则

$$\sqrt{n}\,\frac{\overline{X}-\mu}{\sigma} : \ N(0,1)$$

$$\sum_{i=1}^{n}\frac{(X_i-\overline{X})^2}{\sigma^2} : \ \chi^2(n-1)$$

上面两个随机变量相互独立

$$\frac{\sqrt{n}(\overline{X}-\mu)}{S} : \ t(n-1)$$

虽然这里的公式和课本中“正态总体的样本均值与样本方差”的公式是等价的，但是我觉得我这个版本的公式有几个好处：

1） 公式和公式的证明过程比较接近。因此，内在逻辑更加清楚，至少在我看来更加容易记忆。而且顺便把这两个定理的证明也一块记忆了。

2） 我搞不懂为什么课本要在分母里除以 $\sqrt{n}$，用分子乘以 $\sqrt{n}$ 不是更简洁？

3） 我比较反感 $(n-1)S^2$ 这种写法。S^2 本来就是 $\sum_{i=1}^{n}(X_i-\overline{X})^2$ 除以(n-1)的结果。实际统计计算时肯定是先算 $\sum_{i=1}^{n}(X_i-\overline{X})^2$，所以绝对不会有先将它除以(n-1)得到 S^2，再重新乘回去(n-1)的可笑做法。而且，(n-1)在有些地方作为除数，有些地方作为乘数，学生容易在记忆时发生混淆。

第七章

第七章的公式好像都是理解了就能记住的。所以这里先空着。