[关闭]
@catscarf 2018-01-22 00:04 字数 4573 阅读 7409

概率论与数理统计 公式大全

作者：catpub 新浪微博：@catpub

部分平台可能无法显示公式，若公式显示不正常可以前往知乎或作业部落进行查看

点击前往知乎查看目录与导航

一 概率论的基本概念

- 加法公式
  - $$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{1\le i<j\le n} P(A_iA_j) + \sum_{1\le i<j<k\le n} P(A_iA_jA_k) + \ldots + (-1)^{n-1}P(A_1A_2\ldots A_n)$$
- 条件概率
  - $$P(B|A) = \frac{P(AB)}{P(A)}$$
- 全概率公式
  - $$P(A) = \sum_{j=1}^{n} P(B_j)P(A|B_j)$$
- 贝叶斯公式
  - $$P(B_i|A) = \frac{P(AB_i)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n} P(B_j)P(A|B_j)}$$

二 随机变量及其概率分布随机变量及其概率分布

- 概率密度
  - $$F(x) = \int_{-\infty}^{x} f(t)dt$$
- 常见分布及期望和方差
  - $0-1$ 分布

$$P(X=k) = p^k(1-p)^{n-k}$$

- $$X \sim 0-1(p), E(X)=p, D(X)=p(1-p)$$
- 泊松分布
$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}\ (k=0,1,2,\dots)$$
- $$X \sim \pi(\lambda), E(X)=\lambda, D(X)=\lambda$$
- 正态分布
$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
- $$X \sim N(\mu,\sigma^2), E(X)=\mu, D(X)=\sigma^2$$
- 指数分布
$$f(x)=\lambda e^{-\lambda x}\ x>0$$
- $$X \sim E(\lambda), E(X)=\frac{1}{\mu}, D(X)=\frac{1}{\lambda^2}$$
- 二项分布
$$P(X=k)=C_n^k \cdot P^k \cdot (1-p)^{n-k}$$
- $$X \sim B(n,p), E(X)=np, D(X)=np(1-p)$$
- 均匀分布
$$f(x)=\frac{1}{b-a}\ a \le x < b$$
- $$X \sim U(a,b), E(X)=\frac{a+b}{2}, D(X)=\frac{(b-a)^2}{12}$$

* 随机变量函数的概率密度
  - 若 $Y=g(x)$，$g'(x)>0$ 或 $g'(x)<0$
  - $$f_Y(y)=f_x(h(y)) \cdot |h'(y)|\ \alpha < y < \beta$$
  - $h(y)$ 是 $g(x)$ 的反函数

三 二元随机变量

* 当 X 与 Y 相互独立时，若 $Z=X+Y$
  - $$f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy$$
* 二元连续型随机变量的条件概率密度

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

- $Z = X + Y$的分布
  - $$F_Z(z) = P(Z \le z) = \iint_{x+y\le z} f(x,y)dxdy$$
  - $$f_Z(z) = \int_{-\infty}^{+\infty} f(z-y,y)dy$$
  - 正态分布
  - $$c_0 + c_1X_1 + c_2X_2 + \ldots + c_nX_n \sim N(c_0 + c_1\mu_1 + \ldots + c_n\mu_n, c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + \ldots + c_n^2\sigma_n^2)$$
  - 二项分布
  - $$X + Y \sim B(n_1 + n_2, p)$$
  - 泊松分布
  - $$X + Y \sim \pi(\lambda_1 + \lambda_2)$$
- $max(X,Y)$和$min(X,Y)$的分布
  - $$f_{max}(z) = f_X(z)f_Y(z)$$
  - $$f_{min}(z) = 1 - (1 - f_X(z))(1 - f_Y(z))$$

四 期望与方差

- 数学期望
  - $$E(X) = \sum_{k=1}^{+\infty} x_k p_k$$
  - $$E(X) = \sum_{k=1}^{+\infty} x_k p_k$$
  - $$E(Y) = E[g(X)] = \sum_{k=1}^{\infty} g(x_k)p_k$$
  - $$E(Y) = E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$$
  - $$E(Z) = E[h(X,Y)] = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} h(x_i,y_j)p_{ij}$$
  -

$$E(Z)=E[h(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}h(x,y)f(x,y)dxdy$$

  - $E(cX)=xE(X)$
  - $E(XY)=E(X)E(Y)$ （要求独立）
- 方差
  - $$D(X)=\sum_{i=1}^{+\infty}[x_i-E(X)]^2p_2$$
  - $$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2f(x)dx$$
  - $$D(X)=Var(X)=E\{[X-E(X)^2]\}=E(X^2)-[E(X)]^2$$
  - $D(X+Y)=D(X)+D(Y)+2\cdot Cov$
  - $$E(X^*)=0$$
  - $$D(X^*)=1$$
- 协方差
  - $$Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$$
  - $$Cov(X,Y)=E(XY)-E(X)E(Y)$$
  - $$Cov(aX,bY)=ab\cdot Cov(X,Y)$$
  - $$Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$$
  - $$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}$$
  - $$\rho_{XY}=Cov(X^*,Y^*)$$
- 不相关与独立
  - 不相关：$\rho_{XY}=0$
  - 独立：

$$P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j),\quad f(x,y)=f_X(x)f_Y(y)$$

- 常见分布及期望和方差
  - $0-1$ 分布

$$P(X=k)=p^k(1-p)^{n-k}$$

  - $$X\sim 0-1(p),E(X)=p,D(X)=p(1-p)$$
  - 泊松分布

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

- $$X\sim \pi(\lambda),E(X)=\lambda,D(X)=\lambda$$
- 正态分布

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

- $$X\sim N(\mu,\sigma^2),E(X)=\mu,D(X)=\sigma^2$$
- 指数分布

$$f(x)=\lambda e^{-\lambda x}\ x>0$$

- $$X\sim E(\lambda),E(X)=\frac{1}{\mu},D(X)=\frac{1}{\lambda^2}$$
- 二项分布

$$P(X=k)=C_n^k\cdot P^k\cdot(1-p)^{n-k}$$

- $$X\sim B(n,p),E(X)=np,D(X)=np(1-p)$$
- 均匀分布

$$f(x)=\frac{1}{b-a}\ a\leq x<b$$

- $$X\sim U(a,b),E(X)=\frac{a+b}{2},D(X)=\frac{(b-a)^2}{12}$$

五 切比雪夫不等式，大数定律中心极限定理

- 切比雪夫不等式 Chebyshev's inequality
  - $$P\{|X-\mu|\geq\epsilon\}<\frac{\sigma^2}{\epsilon^2}$$
- 伯努利大数定律
  - $$\lim_{n\to+\infty}P\{|\frac{n_A}{n}-p|\geq\epsilon\}=0$$
- 独立同分布的中心极限定理（CLT）
  - 设 $X_1,X_2,\dots,X_n,\dots$ 相互独立且同分布，$E(X_i)=\mu,D(X_i)=\sigma^2,i=1,2,\dots$ 则对于充分大 n 的，有
  - $$\sum_{i=1}^{n}X_i\sim^{\text{近似}}N(n\mu,n\sigma^2)$$

- 德莫弗-拉普拉斯定理
  - 即二项分布可以用正态分布逼近
  - $$n_A \overset{近似}{\sim} N(np, np(1-p))$$

六 统计量与抽样分布

- 样本均值
  - $$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$
  - $$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$
  - $$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$
- 样本方差
  - $$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$
  - $$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$
  - $E(S^2) = \sigma^2$
  - $$D(S^2) = \frac{2\sigma^4}{n-1}$$
- 样本矩
  - $$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$
  - $$B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k$$
- χ^2分布
  - $$\chi^2 = \sum_{i=1}^{n} X_i^2$$
  - $E(\chi^2) = n$
  - $D(\chi^2) = 2n$
  - $Y_1 + Y_2 \sim \chi^2(n_1+n_2)$
- t分布
  -

$$T = \frac{X}{\sqrt{(Y/n)}}, X \sim N(0,1), Y \sim \chi^2(n)$$

- F 分布
  - $$F = \frac{X/n_1}{Y/n_2}, X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$$
- 两个正态总体的抽样分布
  - $$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2}{S_2^2} \Big/ \frac{\sigma_1^2}{\sigma_2^2} \sim F(n_1-1, n_2-1)$$
  - $$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$
  - $$\frac{(\bar{X}-\bar{Y})-\mu_1-\mu_2}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$$
  - 其中
    $$S_w^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

七 参数估计

- 极大似然估计
  - $$L(\theta) = \prod_{i=1}^{n} p(x_i;\theta)$$
  - $$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$
  - $X \sim N(\mu, \sigma^2), \quad \hat{\mu} = \bar{X}, \quad \hat{\sigma}^2 = B_2$
  - $X \sim U(a,b), \quad \hat{a} = \min\{X_1, \ldots, X_n\}, \quad \hat{b} = \max\{X_1, \ldots, X_n\}$
- 置信区间
  - $P\{\hat{\theta}_L(X_1, \ldots, X_n) < \theta < \hat{\theta}_U(X_1, \ldots, X_2)\} \geq 1-\alpha$

八 假设检验

- 拟合优度检验
  - $$\sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} \sim^{\text{近似}} \chi^2(k-r-1)$$

- n_i 为实际频数，np_i 为理论频数

本章最后修订时间：2018.1.21 如有错误欢迎前往知乎指正

概率论与数理统计 公式大全

作者：catpub 新浪微博：@catpub

部分平台可能无法显示公式，若公式显示不正常可以前往知乎或作业部落进行查看

- **内容目录**

-

-
  -
    - 未分类 14
    - 神经网络和深度学习笔记 第四章 深层神经网络
    - 神经网络和深度学习笔记 第三章 浅层神经网络
    - 神经网络和深度学习笔记 第二章 神经网络基础
    - 吴恩达深度学习系列课程笔记目录与导航
    - 希腊字母表 Markdown版
    - 概率论与数理统计 公式大全
    - 概率论与数理统计 公式大全
    - 概率论与数理统计笔记 第七章 参数估计
    - 概率论与数理统计笔记 第六章 统计量与抽样分布
    - 概率论与数理统计笔记 第五章 大数定律及中心极限定理
    - 概率论与数理统计笔记 第四章 随机变量的数字特征
    - 概率论与数理统计笔记 第三章 二元随机变量及其分布
    - 概率论与数理统计笔记 第二章 随机变量及其概率分布
    - 概率论与数理统计笔记 第一章 概率论的基本概念
  - 搜索 catscarf 的文稿标题，* 显示
  - 以下【标签】将用于标记这篇文稿：

-
-
-

-
  - 下载客户端
  - 关注开发者
  - 报告问题，建议
  - 联系我们
-

添加新批注

保存 取消

在作者公开此批注前，只有你和作者可见。

保存 取消

修改 保存 取消 删除

- 私有
- 公开
- 删除

查看更早的 5 条回复

回复批注

×

通知

取消 确认