概率论与数理逻辑笔记整理V1.0

陈鸿峥

2018.08 *

目录

1 基本概念

1.1 事件与概率

命题 1. 事件的基本运算

1. 分配律：$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$, $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$
2. 德摩根律：$\overline{A\cup B}=\overline{A}\cap\overline{B}$, $\overline{A\cap B}=\overline{A}\cup\overline{B}$

注意概率是对一个集合的函数，有如下定义.

定义 1 (概率)**.** E为随机试验，S是它的样本空间，称$\mathbb{P}(A)$为E中某一事件A的概率，若集合函数$\mathbb{P}(\cdot)$满足：

1. 非负性：对每一个事件A，$\mathbb{P}(A)\geq 0$
2. 规范性：必然事件B，$\mathbb{P}(B)=1$
3. 可列可加性：$A_1,A_2,\ldots$两两不相容，$\mathbb{P}\left(\bigcup\limits_{i=1}^{\infty}A_i\right)=\sum\limits_{i=1}^{\infty}\mathbb{P}(A_i)$

*Build 20180803

由定义可得一些基本性质：

1. $\mathbb{P}(\varnothing)=0, \mathbb{P}(A)\le 1$
2. 有限可加性：$A_1, A_2, \ldots$两两不相容，$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n}\mathbb{P}(A_i)$
3. 若$A\subset B$，则$\mathbb{P}(B-A)=\mathbb{P}(B)-\mathbb{P}(A), \mathbb{P}(B)\ge\mathbb{P}(A)$
4. 逆事件概率：$\mathbb{P}\left(\overline{A}\right)=1-\mathbb{P}(A)$
5. 容斥原理：$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{k=1}^{n}\left((-1)^{k-1}\sum_{\substack{I\subset\{1,\ldots,n\}\\|I|=k}}\mathbb{P}\left(\bigcap_{i\in I} A_i\right)\right)$

定理 1 (实际推断原理). 概率很小的事件在一次试验中实际上几乎是不发生的

1.2 条件概率

定义 2 (条件概率). 设A, B为两个事件，且$\mathbb{P}(A)>0$，则称

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

为在事件A发生的条件下事件B发生的条件概率

定理 2 (乘法公式). 设$A_1, A_2, \ldots, A_n$为n个事件，$n\ge 2$，且$\mathbb{P}(A_1A_2\cdots A_{n-1})>0$，则

$$\mathbb{P}(A_1A_2\cdots A_{n-1})=\mathbb{P}(A_n\mid A_1A_2\cdots A_{n-1})\,\mathbb{P}(A_{n-1}\mid A_1A_2\cdots A_{n-2})\cdots\mathbb{P}(A_2\mid A_1)\,\mathbb{P}(A_1)$$

定义 3 (划分). 两两交为空，所有并为全集

定理 3 (全概率公式). 设试验E的样本空间为S，A为E的事件，$B_1, B_2, \ldots, B_n$为S的一个划分，且$\mathbb{P}(B_i)>0$，则

$$\mathbb{P}(A)=\mathbb{P}(AB_1)+\mathbb{P}(AB_2)+\cdots+\mathbb{P}(AB_n)=\mathbb{P}(A\mid B_1)\,\mathbb{P}(B_1)+\mathbb{P}(A\mid B_2)+\cdots+\mathbb{P}(A\mid B_n)\,\mathbb{P}(B_n)$$

定理 4 (贝叶斯(Bayes)公式). 设试验E的样本空间为S，A为E的事件，$B_1, B_2, \ldots, B_n$为S的一个划分，且$\mathbb{P}(A)>0, \mathbb{P}(B_i)>0$，则

$$\mathbb{P}(B_i\mid A)=\frac{\mathbb{P}(B_iA)}{\mathbb{P}(A)}=\frac{\mathbb{P}(A\mid B_i)\,\mathbb{P}(B_i)}{\sum_{i=1}^{n}P(A\mid B_i)\mathbb{P}(B_i)}$$

特别地，当$n=2$时有

$$\mathbb{P}(B\mid A)=\frac{\mathbb{P}(A\mid B)\,\mathbb{P}(B)}{\mathbb{P}(A\mid B)\,\mathbb{P}(B)+\mathbb{P}\left(A\mid\overline{B}\right)\mathbb{P}\left(\overline{B}\right)}$$

定义 4 (独立性). 对于事件$A_1, \ldots, A_n$，

- 若$\mathbb{P}(A_i\cap A_j)=\mathbb{P}(A_i)\,\mathbb{P}(A_j), \forall i, j$，则称$\{A_1, \ldots, A_n\}$两两(pairwise)独立

- 若$\mathbb{P}\left(\bigcap_{j\in I} A_j\right) = \prod_{j\in I}\mathbb{P}\left(A_j\right), \forall I \in 2^{[n]}$，其中$2^{[n]}$为$\{A_i\}_{i=1}^n$的所有子集，则称$\{A_1,\ldots,A_n\}$相互(mutually)独立

区分以下两个概念

1. A, B对立(exclusive) $\Leftrightarrow \mathbb{P}(A\cap B) = 0$，即不相交(disjoint)
2. A, B独立(independent) $\Leftrightarrow \mathbb{P}(A\cap B) = \mathbb{P}(A)\cdot\mathbb{P}(B)$，即不相关(unrelated)

2 随机变量及其分布

定义 5 (概率质量函数(PMF)). 对于离散随机变量X，$f_X(k) = \mathbb{P}\left(X = k\right)$

定义 6 (概率密度函数(PDF)). 对于连续随机变量X，$f_X(k) = \mathbb{P}\left(X \in A\right) = \int_A f_X(x)dx$

定义 7 (分布函数/累积密度函数(CDF)). $F_X(x) = \mathbb{P}\left(X \le x\right) = \int_{-\infty}^{x} f_X(z)dz$

定义 8 (期望).

$$\mathbb{E}\left(X\right) = \sum_k k f_X(k)$$

$$\mathbb{E}\left(X\right) = \int_x^{-\infty} +\infty \ \mathrm{d}x f_X(x)$$

期望具有线性性，即

$$\mathbb{E}\left(X+Y\right) = \mathbb{E}\left(X\right) + \mathbb{E}\left(Y\right), \mathbb{E}\left(cX\right) = c\mathbb{E}\left(X\right)$$

定义 9 (方差).

$$\mathrm{Var}\left(X\right) = \sigma^2 = \mathbb{E}\left(\left(X - \mathbb{E}\left(X\right)\right)^2\right) = \mathbb{E}\left(X^2\right) - \mathbb{E}\left(X\right)^2 \ge 0$$

由方差定义和期望的线性性有

$$\mathrm{Var}\left(aX+b\right) = a^2\mathrm{Var}\left(X\right)$$

注意方差并不线性

$$\begin{aligned}
\mathrm{Var}\left(X+Y\right) &= \mathbb{E}\left(\left(X+Y\right)^2\right) - \left(\mathbb{E}\left(X\right) + \mathbb{E}\left(Y\right)\right)^2 \\
&= \mathbb{E}\left(X^2\right) - \mathbb{E}\left(X\right)^2 + \mathbb{E}\left(Y^2\right) - \mathbb{E}\left(Y\right)^2 + 2(\mathbb{E}\left(XY\right) - \mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)) \\
&= \mathrm{Var}\left(X\right) + \mathrm{Var}\left(Y\right) + 2\mathrm{Cov}\left(X,Y\right)
\end{aligned}$$

若X, Y独立，则$\mathrm{Cov}\left(X,Y\right) = 0$

定理 5. 若$X_1,\ldots,X_n$独立，则

$$\mathbb{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathbb{E}\left(X_i\right)$$
$$\operatorname{Var}\left(\prod_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}\left(X_i\right)$$
$$\operatorname{Cov}\left(X_i, X_j\right) = 0,\, i \neq j$$

2.1 常见的离散分布

1. 伯努利分布 **Bernoulli(p)**（二项分布的特殊情形）

$$f_X(k) = \begin{cases} 1-p & k=0 \\ p & k=1 \end{cases}$$
$$\mathbb{E}(X) = p$$

2. 二项分布 **Binomial(n,p)**

$$f_X(k) = \binom{n}{k} p^k (1-p)^n - k,\, 0 \leq k \leq n$$
$$\mathbb{E}(X) = n \cdot p$$

e.g. 扔n次硬币扔到k次正面（做实验n次，记录成功的次数）

3. 几何分布 **Geometric(p)**（负二项分布的特殊情形）

$$f_X(k) = (1-p)^k \cdot p, k \geq 0$$
$$\mathbb{E}(X) = \frac{1-p}{p}$$

e.g. 扔k次反面直至扔到正面（做实验直到你成功，记录失败的次数）

4. 负二项分布 **NegetiveBinomial(t,p)**

$$f_X(k) = \binom{k+t-1}{t-1} p^t (1-p)^k,\, k \geq 0$$
$$\mathbb{E}(X) = t \cdot \frac{1-p}{p}$$

e.g. 扔k次反面直到有t个正面（做实验直到你获得t次成功，记录失败次数）

5. 超几何分布 **HyperGeometric(N,n,M)**

$$f_X(k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$
$$\mathbb{E}(X) = n\frac{M}{N}$$

e.g. M个产品中有N个次品，检查n次得到k个次品

6. 泊松分布 **Poisson(λ)**

$$f_X(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$
$$\mathbb{E}(X) = \lambda$$

$X \sim B(n,p)$，若$p = \dfrac{\lambda}{n}$，且n非常大，则

$$\begin{aligned}\mathbb{P}(X = k) &= \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!}\frac{\lambda^k}{n^k}\left(1 + \frac{-\lambda}{n}\right)^{n-k} \\ &\approx \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}\end{aligned}$$

注意泊松分布具有无记忆性(memoryless)，即

$$\mathbb{P}(X \geq a \mid X \geq b) = \mathbb{P}(X \geq a - b)$$

2.2 常见的连续分布

1. 指数分布 **Exponential(λ)**

$$f_X(x) = \lambda \mathrm{e}^{-\lambda x},\ x \geq 0$$
$$\mathbb{E}(()X) = \frac{1}{\lambda}$$

2. 正态分布 **Normal(μ, σ)**

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{\frac{-(x-\mu)^2}{2\sigma^2}}$$
$$\mathbb{E}(X) = \mu$$

算平方

$$
\begin{aligned}
&\left(\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{-x^2}{2}}\,\mathrm{d}x\right)^2\\
&=\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{-x^2}{2}}\,\mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{e}^{\frac{-x^2}{2}}\,\mathrm{d}y\\
&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{-(x^2+y^2)}{2}}\,\mathrm{d}x\,\mathrm{d}y\\
&=\int_{\theta}^{0} 2\pi\ \mathrm{d}\int_{r}^{0} +\infty\ \mathrm{d}r\mathrm{e}^{\frac{-r^2}{2}}\\
&=\int_{\theta}^{0} 2\pi\ \mathrm{d}\int_{u}^{0} +\infty\ \mathrm{d}\mathrm{e}^{-u}\\
&=\int_{\theta}^{0} 2\pi\ \mathrm{d}1\\
&=2\pi
\end{aligned}
$$

3 大数定律

定理 6 (切比雪夫(Chebyshev)不等式)**.**

$$\mathbb{P}\left(|X-\mathbb{E}\left(X\right)|\geq\lambda\right)\leq\frac{\mathbb{E}\left(|X-\mathbb{E}\left(X\right)|\right)}{\lambda}$$

定理 7 (弱大数定律)**.** *若$X_1, X_2, \ldots$为独立随机变量且同等分布，$\mathbb{E}\left(X_i\right)=\mu, \mathrm{Var}\left(X_i\right)=\sigma^2$，则*

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|>\varepsilon\right)=0, \forall\varepsilon>0$$

定理 8 (强大数定律)**.** *若$X_1, X_2, \ldots$为独立随机变量且同等分布，$\mathbb{E}\left(X_i\right)=\mu, \mathrm{Var}\left(X_i\right)=\sigma^2$，则*

$$\mathbb{P}\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}X_i=\mu\right)=1$$

4 统计

定义 10 (估计)**.** *$X_1, X_2, \ldots, X_n$为独立随机变量，从有参数$\mu, \sigma, \theta, \ldots$的分布$f$中得到，对参数$\theta$的估计是函数$T(X_1, \ldots, X_n)$，称$T$是期望(expected)估计，若*

$$\mathbb{E}\left(T(X_1, \ldots, X_n)\right)=\theta$$

合适(probable)的估计，若

$$\mathbb{P}\left(|T(X_1, \ldots, X_n)-\theta|>\varepsilon\right)\to 0,\, n\to\infty$$