AI 中的数学

第五、六讲

方聪，概率统计部分参考章复熹和张原老师课件

2024 年秋季

1. 方差
2. 随机变量的其他数学特征

- 定义 7.1 & 7.2.. 假设 EX 存在, 且 $E(X-EX)^2$ 也存在. 则称 $E(X-EX)^2$ 为 X 的方差, 记为 $\mathrm{var}(X)$ 或 $D(X)$. 称 $\sqrt{\mathrm{var}(X)}$ 为标准差.
- 定理 7.1 (切比雪夫不等式) 假设 $\mathrm{var}(X)$ 都存在, 则 $\forall \varepsilon > 0$, 有
$$P(|X-EX| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^2}\mathrm{var}(X).$$
- 证: $\{|X-EX| \geqslant \varepsilon\} = \left\{(X-EX)^2 \geqslant \varepsilon^2\right\}$, 对 $Y=(X-EX)^2$ 用马尔可夫不等式.
- 推论 7.1. 若 $\mathrm{var}(X)=0$, 则 X 退化.

推论 **7.1**：若 $\mathrm{var}(X)=0$，则

$$P(X=E(X))=1.$$

证明：由切比雪夫不等式知

$$P(|X-E(X)|\geqslant \frac{1}{n})=0 \quad (n=1,2,\cdots),$$

于是

$$\begin{aligned} P(X\neq E(X)) &= P(\bigcup_{n=1}^{\infty}\left\{|X-E(X)|\geqslant \frac{1}{n}\right\}) \\ &\leqslant \sum_{n=1}^{\infty} P(|X-E(X)|\geqslant \frac{1}{n})=0. \end{aligned}$$

所以 $P(X=E(X))=1$。

- 定理 7.2. $\mathrm{var}(X) = EX^2 - (EX)^2$.
- 证:

$$\mathrm{var}(X) = E\left(X^2 - 2X \cdot EX + (EX)^2\right) = EX^2 - (EX)^2$$

- 具体地，离散型或连续型的公式如下:

$$\mathrm{var}(X) = \sum_k x_k^2 p_k - (EX)^2$$

$$\mathrm{var}(X) = \int_{-\infty}^{\infty} x^2 p(x) dx - (EX)^2$$

- X 的线性变换的方差:

$$\mathrm{var}(aX + b) = a^2 \,\mathrm{var}(X)$$

定理：设 X 为随机变量，则方差 $D(X) = \inf_{c\in\mathbb{R}} E(X-c)^2$。

定理：设 X 为随机变量，则方差 $D(X)=\inf_{c\in\mathbb{R}} E(X-c)^2$。

证明：【方法一】利用 $E(X+c)=E(X)+c$ 与 $D(X+c)=D(X)$，可得

$$D(X)=D(X-c)=E(X-c)^2-(E(X-c))^2\leqslant E(X-c)^2.$$

等号成立条件是 $E(X)=c$。
【方法二】利用

$$\begin{aligned}E(X-c)^2&=E(X-E(X)+E(X)-c)^2\\&=D(X)+2E[(X-E(X))(E(X)-c)]+(E(X)-c)^2\\&=D(X)+(E(X)-c)^2.\end{aligned}$$

在 $c=E(X)$ 处取得最小值 $D(X)$。

(1) 两点分布.

设随机变量 X 服从两点分布，即 $X \sim B(1, p)$，根据之前的计算 $E(X) = p$，$E(X^2) = 0^2 \cdot P(X = 0) + 1^2 \cdot P(X = 1) = p$，有

$$D(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1 - p)$$

(3) 泊松分布.

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

- $EX=\lambda$, 且 $\forall k\geqslant 1, kp_k=\lambda p_{k-1}$. 因此, $\forall k\geqslant 2$,

$$k(k-1)p_k=\lambda(k-1)p_{k-1}=\lambda^2 p_{k-2}.$$

- 于是, $EX(X-1)=\lambda^2$, 从而

$$\mathrm{var}(X)=EX^2-(EX)^2=EX(X-1)+EX-(EX)^2=\lambda.$$

(2) 二项分布.

$$P(X=k)=C_n^k p^k q^{n-k}=:b(n;k),k=0,1,\cdots,n,(q=1-p).$$

- $EX=np$, 且 $\forall 1\leqslant k\leqslant n$,

$$k\cdot b(n;k)=np\cdot b(n-1,k-1).$$

- $\forall 2\leqslant k\leqslant n$,

$$\begin{aligned}k(k-1)\cdot b(n;k)&=np\cdot(k-1)\cdot b(n-1,k-1)\\&=np\cdot(n-1)p\cdot b(n-2,k-2)\end{aligned}$$

- 于是, $EX(X-1)=np(n-1)p=(np)^2-np^2$, 从而

$$D(X)=EX^2-(EX)^2=EX(X-1)+EX-(EX)^2=npq.$$

(4) 均匀分布

设随机变量 X 服从区间 $[a,b]$ 上的均匀分布，即 X 有密度分布：

$$p(x)=\begin{cases}\frac{1}{b-a}, & a\leqslant x\leqslant b,\\ 0, & \text{其他}\end{cases}$$

已经计算 X 的期望 $E(x)=\frac{a+b}{2}$，且

$$\int_{-\infty}^{+\infty}x^2p(x)dx=\frac{1}{b-a}\int_a^b x^2dx=\frac{b^2+ab+a^2}{3}.$$

则有

$$D(X)=\frac{b^2+ab+a^2}{3}-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}.$$

(5) 指数分布

设随机变量 X 服从参数为 λ 的指数分布，即 X 的分布函数为:

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & X \leqslant 0, \end{cases}$$

已经计算期望 $E(X) = \frac{1}{\lambda}$，且

$$\int_{-\infty}^{+\infty} x^2 p(x) dx = \lambda \int_0^{+\infty} x^2 e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

则有

$$D(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

(6) 正态分布.

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

- 若 $\mu = EX = 0, \sigma^2 = 1$, 则,

$$\begin{aligned}\mathrm{var}(X) = EX^2 &= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= -\frac{2}{\sqrt{2\pi}} \int_0^{\infty} x d e^{-\frac{x^2}{2}} = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{x^2}{2}} dx = 1\end{aligned}$$

- 一般情形, $Y = \frac{X-\mu}{\sigma} \sim N(0,1)$. 则

$$\begin{aligned}&X - EX = (\mu + \sigma Y) - (\mu + \sigma EY) = \sigma(Y - EY) \\ &\Rightarrow \mathrm{var}(X) = E(\sigma(Y - EY))^2 = \sigma^2 \,\mathrm{var}(Y) = \sigma^2\end{aligned}$$

(7) 伽马分布

设随机变量 X 服从伽马分布，有分布密度

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0$$

已经计算 X 的期望 $E(X) = \frac{\alpha}{\beta}$，则有

$$D(X) = \int_0^{+\infty} x^2 \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx - \left(\frac{\alpha}{\beta}\right)^2$$

做变量替换 $\beta x = t$，易知

$$\begin{aligned} D(X) &= \frac{1}{\Gamma(\alpha)\beta^2} \int_0^{+\infty} t^{\alpha+1} e^{-t} dt - \left(\frac{\alpha}{\beta}\right)^2 \\ &= \frac{1}{\Gamma(\alpha)\beta^2} \Gamma(\alpha+2) - \left(\frac{\alpha}{\beta}\right)^2 \\ &= \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)\beta^2} - \left(\frac{\alpha}{\beta}\right)^2 = \frac{\alpha}{\beta^2} \end{aligned}$$

随机变量的标准化：一般地，若 X 的方差存在，且 $\operatorname{var}(X)>0$，则

$$X^{*}=\frac{X-E(X)}{\sqrt{\operatorname{var}(X)}}$$

满足 $E(X^{*})=0, \operatorname{var}\left(X^{*}\right)=1$. 称 X^{*} 为 X 的标准化。

例：设随机变量 $X \sim N(0,1)$，则对一切正整数 k，

$$E(X^{2k-1}) = 0, \quad E(X^{2k}) = (2k-1)(2k-3)\cdots\cdot 3\cdot 1.$$

例：设随机变量 $X \sim N(0,1)$，则对一切正整数 k，

$$E(X^{2k-1}) = 0, \quad E(X^{2k}) = (2k-1)(2k-3)\cdots 3\cdot 1.$$

证明：对任何 $m \geqslant 1$，积分 $\int_{-\infty}^{+\infty} |x|^m e^{-x^2/2} dx$ 收敛，因此

$E(X^m)$ 存在，由于

$$x^{2k-1}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

是 x 的奇函数，故

$$E(X^{2k-1}) = \int_{-\infty}^{+\infty} x^{2k-1}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx = 0$$

$$
\begin{aligned}
E(X^{2k}) &= \int_{-\infty}^{+\infty} x^{2k} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^{2k-1} d(e^{-x^2/2}) \\
&= (2k-1) \int_{-\infty}^{+\infty} x^{2k-2} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\
&= (2k-1) E(X^{2k-2}).
\end{aligned}
$$

这是递推公式，故

$$
E(X^{2k}) = (2k-1)(2k-3) \cdots \cdot 3 \cdot 1
$$

例：设随机变量 X 的概率密度函数为 $f(x)=\frac{k\theta^k}{x^{k+1}}, x>\theta>0$, $k>2$ 为正整数，求 (1)$E(X)$, (2)$D(X)$.

例：设随机变量 X 的概率密度函数为 $f(x) = \frac{k\theta^k}{x^{k+1}}, x > \theta > 0$, $k > 2$ 为正整数，求 (1)$E(X)$, (2)$D(X)$.

解：(1):

$$E(X) = \int_\theta^{+\infty} \frac{k\theta^k}{x^{k+1}} x dx = \frac{k\theta}{k-1}.$$

(2):

$$D(X) = \int_\theta^{+\infty} \frac{k\theta^k}{x^{k+1}} x^2 dx - (E(X))^2 = \frac{3k-2k^2}{(k-2)(k-1)^2}.$$

例：设连续随机变量 X 的分布函数为 $F(x)$，且数学期望存在，证明：

$$E(X)=\int_0^{\infty}[1-F(x)]dx-\int_{-\infty}^{0}F(x)dx.$$

例：设连续随机变量 X 的分布函数为 $F(x)$，且数学期望存在，证明：

$$E(X) = \int_0^{\infty} [1 - F(x)] dx - \int_{-\infty}^{0} F(x) dx.$$

证明：

$$E(X) = \int_{-\infty}^{\infty} xp(x)dx = \int_{-\infty}^{0} xp(x)dx + \int_{0}^{\infty} xp(x)dx.$$

将第一个积分改写为：

$$\begin{aligned} \int_{-\infty}^{0} xp(x)dx &= \int_{-\infty}^{0} \left(\int_{x}^{0} dy \right) p(x)dx \\ &= -\int_{-\infty}^{0} \int_{-\infty}^{y} p(x)dxdy \\ &= -\int_{-\infty}^{0} F(y)dy. \end{aligned}$$

第二个积分同理，

$$\begin{aligned}\int_0^{\infty} xp(x)dx &= \int_0^{\infty}\left(\int_0^{x} dy\right)p(x)dx \\ &= \int_0^{\infty}\int_y^{\infty} p(x)dxdy \\ &= \int_0^{\infty}[1-F(y)]dy.\end{aligned}$$

将二式加和即可得

$$E(X)=\int_0^{\infty}[1-F(x)]dx-\int_{-\infty}^{0}F(x)dx.$$

1 方差

2 随机变量的其他数学特征

设 X 是随机变量，如果 $E(X^k)$ 存在（k 是正整数），则称 $E(X^k)$ 是 X 的 k 阶原点矩，常常记为 ν_k。

设 X 是随机变量，如果 $E(X)$ 存在，且 $E(X-E(X))^k$ 存在（k 是正整数），则称 $E(X-E(X))^k$ 为 X 的 k 阶中心矩，常常记为 μ_k。

显然，$E(X)=\nu_1$，$\mathrm{var}(X)=\mu_2$。

随机变量的 p 分位数：若 X 是随机变量，$0 < p < 1$，且

$$P(X < a) \leqslant p \leqslant P(X \leqslant a),$$

则称 a 为 X 的一个 p 分位数。

$p = 0.5$ 时，也称 a 为一个中位数.

例：设随机变量 X 的可能值是 $1,2,3$ 且

$$P(X=1)=\frac{1}{3},\quad P(X=2)=\frac{1}{6},\quad P(X=3)=\frac{1}{2}.$$

则 $E(X)=\frac{13}{6}$，中位数有无穷个，区间 $[2,3]$ 中的每个数都是 X 的中位数。

定理：设 $X = X(\omega)$ 是随机变量，对某个 $\alpha \geqslant 1$，$E(|X|^{\alpha})$ 存在，则 $E(X)$ 存在，且

$$E(|X|) \leqslant (E(|X|^{\alpha}))^{1/\alpha}.$$

定理：设 $X=X(\omega)$ 是随机变量，对某个 $\alpha\geqslant 1$，$E(|X|^\alpha)$ 存在，则 $E(X)$ 存在，且

$$E(|X|)\leqslant (E(|X|^\alpha))^{1/\alpha}.$$

证明：首先指出，对一切 $x\geqslant 0$，$\alpha\geqslant 0$，如下不等式成立：

$$x^\alpha\geqslant a^\alpha+\alpha a^{\alpha-1}(x-a). \tag{1}$$

实际上，令 $f(x)=x^\alpha-a^\alpha-\alpha a^{\alpha-1}(x-a)$，则
$f'(x)=\alpha(x^{\alpha-1}-a^{\alpha-1})$，从而 $f(x)$ 在 $x=a$ 处达到最小值，由于 $f(a)=0$，因此式1成立。
由于 $\alpha\geqslant 1$，有 $|X(\omega)|\leqslant |X(\omega)|^\alpha+1$，知 $E(X)$ 存在。令 $a=E(X)$，由式(1)知

$$|X(\omega)|^\alpha\geqslant (E(|X|))^\alpha+\alpha(E(|X|))^{\alpha-1}(|X(\omega)|-E(|X|)).$$

两侧取数学期望，得 $(E(|X|))^\alpha\leqslant E(|X|^\alpha)$，表明定理成立。

例：设 X 是仅取非负整数的离散随机变量，若其数学期望存在，证明

(1)$E(X) = \sum_{k=1}^{\infty} P(X \geqslant k)$.

(2)$\sum_{k=0}^{\infty} kP(X > k) = \frac{1}{2}[E(X^2) - E(X)]$.

例：设 X 是仅取非负整数的离散随机变量，若其数学期望存在，证明

(1)$E(X)=\sum_{k=1}^{\infty}P(X\geqslant k)$.

(2)$\sum_{k=0}^{\infty}kP(X>k)=\frac{1}{2}[E(X^2)-E(X)]$.

证明：(1) 由于 $E(X)=\sum_{k=1}^{\infty}kP(X=k)$ 存在，所以该级数绝对收敛，从而

$$E(X)=\sum_{k=1}^{\infty}kP(X=k)=\sum_{k=1}^{\infty}\left[\sum_{i=1}^{k}P(X=k)\right]$$
$$=\sum_{i=1}^{\infty}\left[\sum_{k=i}^{\infty}P(X=k)\right]=\sum_{k=1}^{\infty}P(X\geqslant k).$$

(2)

$$
\begin{aligned}
\sum_{k=0}^{\infty} kP(X>k) &= \sum_{k=0}^{\infty} k \sum_{i=k+1}^{\infty} P(X=i) = \sum_{i=1}^{\infty} \sum_{k=0}^{i-1} kP(X=i) \\
&= \sum_{i=1}^{\infty} P(X=i) \frac{(i-1)i}{2} \\
&= \frac{1}{2} \sum_{i=1}^{\infty} i^2 P(X=i) - \frac{1}{2} \sum_{i=1}^{\infty} iP(X=i) \\
&= \frac{1}{2} E(X^2) - \frac{1}{2} E(X).
\end{aligned}
$$

例：甲乙两人进行象棋比赛，每局甲胜的概率为 p，乙胜的概率为 $q = 1 - p$，比赛进行到有一人连胜两局为止，求平均比赛局数。

例：甲乙两人进行象棋比赛，每局甲胜的概率为 p，乙胜的概率为 $q=1-p$，比赛进行到有一人连胜两局为止，求平均比赛局数。

解：设 X 为决定胜负所需的局数，可以取值为 $2,3,\cdots$，事件 $\{X \geqslant k\}$ 表示“到 $k-1$ 局时没有一人连胜两局”，所以

$$P(X \geqslant 1)=1,$$

$$P(X \geqslant 2k)=p^k q^{k-1}+p^{k-1}q^k=(pq)^{k-1}, \quad k=1,2,\cdots,$$

$$P(X \geqslant 2k+1)=2p^k q^k, \quad k=1,2,\cdots.$$

利用上一题第一问提供的公式，可得

$$E(X)=\sum_{k=1}^{\infty} P(X \geqslant k)=1+\sum_{k=1}^{\infty}(pq)^{k-1}+2\sum_{k=1}^{\infty}(pq)^k$$

$$=1+\frac{1}{1-pq}+\frac{2pq}{1-pq}=\frac{2+pq}{1-pq}.$$

注意到对任意的 $0 < p < 1$ 总有 $p(1-p) \leqslant \frac{1}{4}$，故由 $E(X)$ 关于 pq 单调增可得

$$E(X) \leqslant \frac{2+\frac{1}{4}}{1-\frac{1}{4}} = 3$$

故这种比赛最终决定胜负的平均局数不超过 3 局，在 $p = \frac{1}{2}$ 时达到上界。