第 4 周讲稿

§ 2 数学期望与中位数

一. 定义与计算

定义(随机变量的数学期望)： 设 X 是概率空间（$\Omega, \mathcal{F}, P$）上的一个随机变量，其概率分布为:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

如果级数 $\sum_n p_n x_n$ 绝对收敛，则 X 的**数学期望**定义为 $\sum_n p_n x_n$，并记为 $E(X(\omega))$ 或简单地记为 EX. 数学期望就是（加权）平均值, 所以在统计学中也称为**均值**.

注：1. 为什么需要绝对收敛？

2. 设 X 的分布为：$P(X=(-1)^k \frac{3^k}{k})=\frac{2}{3^k}$, $k=1,2,\cdots$，问 EX 存在吗？

3. 由于离散随机变量的分布函数为 $F(x)=\sum_{i:x_i\le x} p_i=\sum_{i=1}^{\infty} p_i \varepsilon(x-x_i)$，其中 $\varepsilon(x)=\begin{cases}1, & x\ge 0,\\ 0, & x<0.\end{cases}$，故 $EX=\int_{-\infty}^{\infty} x dF(x)$.

例 1：在一个抛骰子的游戏中，你在每轮抛掷中可以获得所抛出的点数两倍的奖金，那么，为参加这一游戏，你付出的“合理”的价格应该是多少？

显然，$\Omega=\{\omega_1,\omega_2,\cdots,\omega_6\}=\{1,2,\cdots,6\}$，设 X 为你的支出，$X(\omega_i)=2\omega_i$，且具有等可能分布

$$X \sim \begin{pmatrix} 2 & 4 & \cdots & 12 \\ \frac{1}{6} & \frac{1}{6} & \cdots & \frac{1}{6} \end{pmatrix}$$

故 $EX=\sum_{i=1}^{6} 2\omega_i P(X=2\omega_i)=\sum_{i=1}^{6} 2i\frac{1}{6}=7$

答：平均每轮 7 元。

例 2：抛一枚均匀硬币，你期望要抛多少次才出现正面？

显然这是一个几何分布的期望问题，假设 $X\sim Ge(p)$，即 $P(X=k)=q^{k-1}p, k=1,2,\cdots$。

则 $EX=\sum_{k=1}^{\infty} kq^{k-1}p=\frac{1}{p}$。

这里 $p=\frac{1}{2}$，所以，答案为 2 次。

例 3（参考 St. Petersburg 悖论）：D. Bernoulli 1738 年提出加倍策略问题。

在抛均匀硬币的游戏中，你采用赌注加倍的策略会怎样？比如，你押 2 元钱赌第 1 次就出现正面，如果赌输了，就押 4 元钱赌第 2 次出现正面，如此继续，一直用加倍赌注的方法去赌。这时的情况会怎么样呢？

首先，你博弈获胜时所投入的资金的期望值为$EX=\sum_{n=1}^{\infty}2^n\frac{1}{2^n}=+\infty$。

而当你首次赢的时候（比如在第 n 次抛时），你已经损失了$\sum_{k=1}^{n-1}2^k=2^n-2$元钱，但第 n 次抛时你赢了2^n元，因此，不管你抛几次，你的净收益是 2 元。注意到第 n 次抛时你赢的概率为$\frac{1}{2^n}$，而你等到抛出正面时的博弈次数是服从参数为$\frac{1}{2}$的几何分布。如何解释这种现象？这种策略成立的前提是你要有无穷的资金的支持，但如果你是融资得到的资金，这个策略可以证明也是不可行的。

另一种描述：1730 年代，数学家丹尼尔•伯努利的堂兄尼古拉一世•伯努利，在致法国数学家皮耶•黑蒙•德蒙马特的信件中，提出一个问题：掷硬币，若第一次掷出正面，你就赚 1 元。若第一次掷出反面，那就要再掷一次，若第二次掷的是正面，你便赚 2 元。若第二次掷出反面，那就要掷第三次，若第三次掷的是正面，你便赚2^2元，……，如此类推，即可能掷一次游戏便结束，也可能反复掷没完没了。问题是，你最多肯付多少钱参加这个游戏？

你最多肯付的钱应等于该游戏的期望值：即

$$EX=\sum_{n=1}^{\infty}2^{n-1}\frac{1}{2^n}=+\infty$$

这个游戏的期望值是无限大，你甚至肯付出无限的金钱去参加这个游戏。但实际中是没几个人愿意拿出几十元或更多的钱去冒险。

解决方法：期望效用理论。赌徒资金为 x，效用函数为对数效用$U(x)=\ln x$，c 为肯付的门票费，则此彩票的期望效用为有限。

$$EU=\sum_{n=1}^{\infty}\frac{\ln(x+2^{n-1}-c)-\ln x}{2^n}<+\infty$$

（比如：只有 2 元钱最多会付 2 元，有 1000 元最多付 5.94 元）

如果赌场只有有限资金W，这时彩票的期望值为

$$EV=\sum_{n=1}^{\infty}\min(2^{n-1},W)\frac{1}{2^n}=\sum_{n=1}^{L}2^{n-1}\frac{1}{2^n}+\sum_{n=L+1}^{\infty}W\frac{1}{2^n}$$
$$=\frac{L}{2}+\frac{W}{2^L}$$

其中 $L=[\log_2 W]$ 实际为赌场在无力支付时赌博最多可能的局数，如果 $W=10$ 亿美元，则 $EV=15.93$ 美元。如果 $W=10^{100}$ 美元，则 $EV=166.5$ 美元。

例 4： $X\sim B(n,p)\Rightarrow EX=np$ **.**

例 5（截止的几何分布的期望）：某射手每次射击的命中率为 $0<p<1$，各次射击是相互独立的，现在他拿了 m 发子弹，射击进行到击中目标或子弹用完为止，求直到射击结束时他平均射击了几次？

解：

$$EX=\sum_{k=1}^{m}kP(X=k)=\sum_{k=1}^{m-1}kq^{k-1}p+mq^{m-1}$$
$$=p\frac{1-mq^{m-1}+(m-1)q^m}{(1-q)^2}+mq^{m-1}$$
$$=\frac{1-q^m}{p}$$

注：1. $m\to\infty$ 时，$EX\to\dfrac{1}{p}$ 与几何分布一致。

2. 可以利用微积分知识求

$$\sum_{k=1}^{m}kx^{k-1}=(\sum_{k=1}^{m}x^k)'=(\frac{x-x^{m+1}}{1-x})'$$
$$=\frac{1-(m+1)x^{m-1}+mx^m}{(1-x)^2}$$

注：截止的几何分布：

背景：做一系列独立重复的 Bernoulli 试验，直到首次试验成功时或试验进行到第 m 次为止，所需的试验的次数。

$$P(X=k)=\begin{cases}q^{k-1}p, & k=1,2,\cdots,m-1\\ q^{m-1}, & k=m\end{cases}$$

定义(随机变量的中位数)：一个数 x 如果满足

$$P(X\le x)\ge p,P(X\ge x)\ge 1-p,\quad 0<p<1$$

则称 x 为随机变量 X 的 $p-$ 分位数，特别，当 $p=\dfrac{1}{2}$ 时，称之为 X 的**中位数**（median）。

显然，x 为随机变量 X 的中位数 $\Leftrightarrow \dfrac{1}{2} \le F(x) \le \dfrac{1}{2} + P(X = x)$，且如果随机变量 X 的分布是对称的，那么其对称中心显然是它的分布的中位数，记为 x_{med} 。

例 6：随机变量 X 的分布为

$$X \sim \begin{pmatrix} -2 & 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$$

则 $EX = \frac{1}{6}$。其中位数等于多少？由于

$P(X \le 0) = \dfrac{1}{2}, P(X \ge 0) = \dfrac{3}{4} > \dfrac{1}{2}$，**显然 0 是它的一个中位数，而当 $0 < x < 1$ 时，**

$P(X \le x) = P(X = -2) + P(X = 0) = \dfrac{1}{2}$；$P(X \ge x) = P(X = 1) + P(X = 2) = \dfrac{1}{2}$。

故每个 $x, 0 \le x < 1$，都是 X 的中位数（有无穷个）。

- **随机变量的众数**（mode）m_z ：随机变量取最大概率的值 m_z ，即

  $P(X = m_z) \ge P(X = x), \forall x$

注：连续随机变量时为密度函数的最大值点。同中位数一样，众数也可能不唯一。

补充定理：$\min\limits_{C} E|X - C| = E|X - x_{med}|$.

证明：只需证明 $E|X - x_{med}| \le E|X - C|, \forall C$

不妨假设 $x_{med} < C$**，若** $X \le x_{med}$ **，则**

$Y \triangleq |X - C| - |X - x_{med}| = C - X - (x_{med} - X) = C - x_{med}$,

而若 $X > x_{med}$**，则**

$Y \triangleq |X - C| - |X - x_{med}| \ge X - C - (X - x_{med}) = x_{med} - C$,

故

$$\begin{aligned} EY &= E[Y 1_{\{X \le x_{med}\}}] + E[Y 1_{\{X > x_{med}\}}] \\ &\ge (C - x_{med}) E 1_{\{X \le x_{med}\}} + (x_{med} - C) E 1_{\{X > x_{med}\}} \\ &= (C - x_{med}) P(X \le x_{med}) + (x_{med} - C) P(X > x_{med}) \\ &= (C - x_{med})[2P(X \le x_{med}) - 1] \end{aligned}$$

由中位数的定义，有 $2P(X \le x_{med}) - 1 \ge 0$**，故** $EY \ge 0$ **，得证。**

二. 实用统计学定律

定理 1（实用统计学定律）：设随机变量 X 的分布函数为 $F_X(x)$，$g(x)$ 为 R 上的连续函数，如果 $g(X)$ 的期望存在，则 $Eg(X)=\int_{-\infty}^{\infty}g(x)dF_X(x)$，特别，当 X 为离散型随机变量时，有 $Eg(X)=\sum_{i=1}^{\infty}g(x_i)p_i$.

证明：不妨设 X 为离散型随机变量，其分布列为 p_i，记 $Y=g(X)$，显然它也是一个离散型随机变量，其取值为 $y_j\in g(X)$ 的值域，且具有分布

$$P(Y=y_j)=P(g(X)=y_j)=P(\bigcup_a[X=a,g(a)=y_j])=\sum_a P(X=a,g(a)=y_j),$$

因此，Y 的期望为

$$EY=\sum_{j=1}^{\infty}y_jP(Y=y_j)$$

$$=\sum_{j=1}^{\infty}y_j\sum_a P(X=a,g(a)=y_j)=\sum_a\sum_{j=1}^{\infty}g(a)P(X=a,g(a)=y_j)=\sum_a g(a)P(X=a)$$

即 $Eg(X)=\sum_{i=1}^{\infty}g(x_i)p_i$。

课上：$P(Y=y_j)=\sum_{i:g(x_i)=y_j}p_i$，

$$EY=\sum_{j=1}^{\infty}y_jP(Y=y_j)=\sum_{j=1}^{\infty}y_j\sum_{i:g(x_i)=y_j}p_i=\sum_j\sum_i y_jp_i1_{\{g(x_i)=y_j\}}$$
$$=\sum_i\sum_j g(x_i)p_i1_{\{g(x_i)=y_j\}}=\sum_i g(x_i)p_i\sum_j 1_{\{g(x_i)=y_j\}}=\sum_i g(x_i)p_i$$

定理 2：对任意随机变量 X，有 $EX=\int_0^{\infty}P(X>x)dx-\int_{-\infty}^{0}P(X\le x)dx$（如果后面的两个积分收敛的话）（提示：分部积分）

§ 3 随机向量

多维随机变量的相关问题

★ **问题提出**：(X,Y) **分量之间可能不独立。**

解决方法：考虑 $P(X \in A, Y \in B), \forall A, B \subset R$

1．二维随机变量的联合分布函数与边缘分布函数

★ 对任意实数 x, y，称函数 $F(x,y) = P(X \le x, Y \le y)$ 为**二维随机变量** (X,Y) **的联合分布函数。**

★ **联合分布函数** $F(x,y)$ 具有如下性质：

（1）$F(x,y)$ 是 x 或 y 的不减函数。

（2）$0 \le F(x,y) \le 1$，且 $F(x,-\infty) = F(-\infty,y) = F(-\infty,-\infty) = 0, F(+\infty,+\infty) = 1$。

（3）$F(x,y)$ 关于 x 或 y 是右连续的。

（4）对任意 $x_1 \le x_2, y_1 \le y_2$，$F(x_2,y_2) - F(x_2,y_1) - F(x_1,y_2) + F(x_1,y_1) \ge 0$。

注：二元实函数 $F(x,y)$ 为某一随机向量的分布函数当且仅当性质（1）—（4）成立。

例子：$F(x,y) = \begin{cases} 1, & x+y \ge 1, \\ 0, & x+y < 1, \end{cases}$ 满足前三条，但不满足第 4 条。

★ X 的边缘分布函数为：$F_X(x) = F(x,+\infty)$；

Y 的边缘分布函数为：$F_Y(x) = F(+\infty, y)$。

2. 独立性的判断

★ X 与 Y 相互独立 $\Leftrightarrow P(X \in A, Y \in B) = P(X \in A)P(Y \in B), \forall A, B \subset R$

$$\Leftrightarrow F(x,y) = F_X(x)F_Y(y) \quad \forall x, y$$

$$\Leftrightarrow p_{ij} = p_{i\bullet}p_{\bullet j} \quad \forall i, j$$ **（离散型）（见下文）**

3. 二维离散型随机变量的联合分布律与边缘分布律

联合分布律：$P(X = x_i, Y = y_j) = p_{ij}$，

$\{p_{ij}\}$为联合分布律$\Leftrightarrow\begin{cases} p_{ij}\ge 0, \\ \sum\limits_{i,j} p_{ij}=1 \end{cases}$

边缘分布律：$P(X=x_i)=\sum\limits_j p_{ij}\equiv p_{i\bullet}$，$P(Y=y_j)=\sum\limits_i p_{ij}\equiv p_{\bullet j}$

离散型条件分布律：在$\{Y=y_j\}$**的条件下，**X**的条件分布律为**

$$P(X=x_i\mid Y=y_j)=\frac{p_{ij}}{p_{\bullet j}} \qquad (i=1,2,\cdots,p_{\bullet j}>0)$$

边缘分布律、条件分布律的例子

例 1 掷两颗均匀骰子，记第一颗骰子出现的点数为 X，而两颗骰子中点数的最大值为 Y，求(X,Y)的联合分布律.

Y \\ X	1	2	3	4	5	6	$p_{i\cdot}$
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	0	2/36	1/36	1/36	1/36	1/36	1/6
3	0	0	3/36	1/36	1/36	1/36	1/6
4	0	0	0	4/36	1/36	1/36	1/6
5	0	0	0	0	5/36	1/36	1/6
6	0	0	0	0	0	6/36	1/6
$p_{\cdot j}$	1/36	3/36	5/36	7/36	9/36	11/36	1

X 与 Y 不相互独立

例 2 在$\{1,2,3,4\}$中任取一数，记为 X，再从$\{1,2,\cdots,X\}$中任取一数，记为 Y，求(X,Y)的联合分布律以及关于 Y 的边缘分布。

$$P(X=i,Y=j)=P(X=i)P(Y=j\mid X=i)=\begin{cases} \frac{1}{4}\times\frac{1}{i}, & 1\le j\le i, \\ 0, & j>i. \end{cases}$$

Y \\ X	1	2	3	4	$p_{i\cdot}$
1	1/4	0	0	0	1/4
2	1/8	1/8	0	0	1/4
3	1/12	1/12	1/12	0	1/4
4	1/16	1/16	1/16	1/16	1/4
$p_{\cdot j}$	25/48	13/48	5/36	7/36	1

4. 离散随机变（向）量的函数的分布律

一维情形：

$$X \sim \{p_i\} \Rightarrow P[g(X)=y]=\sum_{i:g(x_i)=y} p_i, y\in\{g(x_1),\cdots,g(x_i),\cdots\}$$

二维情形：

$$(X,Y) \sim \{p_{ij}\} \Rightarrow P[g(X,Y)=y]=\sum_{i,j:g(x_i,y_j)=y} p_{ij}, y\in\{g(x_1,y_1),\cdots,g(x_i,y_j),\cdots\}$$

★ 列表法：用例子说明。

5. 随机变量函数的数学期望的计算公式

★ $E[g(X)]=\int_{-\infty}^{\infty} g(x)dF(x)=\sum_i g(x_i)p_i,\qquad X\sim\{p_i\}$（后者绝对收敛）

★ $E[g(X,Y)]=\sum_i\sum_j g(x_i,y_j)p_{ij},\qquad (X,Y)\sim\{p_{ij}\}$,（后者绝对收敛）

★ k 阶（原点）矩：$E(X^k)$；k 阶（中心）矩：$E((X-EX)^k)$

X 和 Y 的 $k+l$ 阶混合矩：$E(X^kY^l)$；$k+l$ 阶混合中心矩：$E((X-EX)^k(Y-EY)^l)$

高阶矩存在，则低阶矩必存在（如何证明？）

6. 数学期望与方差的性质

★ 数学期望的性质

性质 1： $|E(X)|\le E(|X|)$.

性质 2： 如果存在 a, b，使得 $P(a\le X\le b)=1$，则 $E(X)$ 存在，而且 $a\le E(X)\le b$.

性质 3： 设 $X_1, X_2,\cdots X_n$ 是 n 个数学期望有意义的随机变量，$c_0,c_1,\cdots,c_n$ 是 n+1 个实数，则 $E(c_0+c_1X_1+\cdots+c_nX_n)=c_0+c_1E(X_1)+\cdots+c_nE(X_n)$.

性质 3： 设 $X_1, X_2,\cdots X_n$ 相互独立，则 $E(\prod_{i=1}^n X_i)=\prod_{i=1}^n EX_i$.

性质 5：(Cauchy-Schwartz 不等式) 若 $E(X^2)<+\infty$，$E(Y^2)<+\infty$，则

$$(E(XY))^2\le E(X^2)E(Y^2).$$

性质 6： $\min_C\{E((X-C)^2)\}=E((X-EX)^2)$.

性质 7： $E(X^2)\ge 0$,且 $E(X^2)=0 \Leftrightarrow P(X=0)=1$。

★随机变量的分解法

例 3（二项分布与超几何分布的均值）；
例 4（匹配数的均值）

★ 方差的性质

性质 1： 若 EX^2 存在，则 $DX = EX^2 - (EX)^2$.

性质 2： $D(C) = 0$,且 $DX = 0 \Leftrightarrow P(X = EX) = 1$.

性质 3： 设 $X_1, X_2, \cdots X_n$ 相互独立，则 $D(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 DX_i$.

补充性质：

- Markov 不等式：$X \ge 0$, 则 $\forall c > 0$，有 $P(X > c) \le \dfrac{EX}{c}$。（提示：$\forall c > 0$，有 $X \ge c1_{\{X>c\}}$）

更一般有：若 X 是实值随机变量，$f: R \to R^+$ 单增，则 $\forall c \in R$，有 $P(X > c) \le \dfrac{Ef(X)}{f(c)}$

- Chebyshev 不等式：$EX^2 < \infty$，则 $\forall \varepsilon > 0$，有 $P(|X - EX| \ge \varepsilon) \le \dfrac{DX}{\varepsilon^2}$。（提示：对 $|X - EX|^2$ 使用 Markov 不等式）。

- 高阶矩存在，则其低阶矩一定存在。

证明：若 $E(X^t)$ 存在，即 $E(|X|^t) = \int_{-\infty}^{\infty} |x|^t f(x)dx < +\infty$ ，故 $\forall s < t$ ，有

$$E(|X|^s) = \int_{|x|^s \le 1} |x|^s \ f(x)dx + \int_{|x|^s > 1} |x|^s \ f(x)dx$$

$$\le \int_{|x|^s \le 1} f(x)dx + \int_{|x|^s > 1} |x|^t \ f(x)dx$$

$$\le P(|X|^s \le 1) + E(|X|^t) < \infty$$

- 若 $E(|X|^k) < +\infty$，则 $\lim_{n\to\infty} n^k P(|X| > n) = 0$ 。

证明：$E(|X|^k) = \int_{-\infty}^{\infty} |x|^k f(x)dx < +\infty$，故 $\lim_{n\to\infty} \int_{|x|>n} |x|^k \ f(x)dx = 0$，

$$n^k P(|X| > n) \le \int_{|x|>n} |x|^k \ f(x)dx \to 0$$

注：反之不成立。

- $EX = \int_0^{\infty} P(X > x)dx - \int_{-\infty}^{0} P(X \le x)dx = \int_0^{\infty} [1 - F(x)]dx - \int_{-\infty}^{0} F(x)dx$
- $\int_{-\infty}^{EX} F(x)dx = \int_{EX}^{\infty} (1 - F(x))dx$