概率统计第五讲：方差

史灵生　清华数学系

1、本讲提要

1. 方差
   - 矩的性质
   - 方差定义
   - 方差性质
2. 常用离散分布
   - 二项分布
   - Poisson分布
   - （超）几何分布

2、矩的性质

性质

$$EX^2 = 0 \Leftrightarrow P(X = 0) = 1.$$

证明：（假设X为离散型）

“$\Leftarrow$” 显然。下证 “$\Rightarrow$”：

- $EX^2 = \sum_i x_i^2 p_i = 0$，
- $p_i > 0 \Rightarrow x_i = 0$，
- $P(X = 0) = 1$。

3、数学期望的统计意义

考虑 $E(X-m)^2 = \min_a E(X-a)^2$，$m=$?

$$\begin{aligned} E(X-a)^2 &= E(X-EX+EX-a)^2 \\ &= E[(X-EX)^2+2(X-EX)(EX-a)+(EX-a)^2] \\ &= E(X-EX)^2+2(EX-a)E(X-EX)+(EX-a)^2 \\ &= E(X-EX)^2+(EX-a)^2 \geq E(X-EX)^2 \end{aligned}$$

所以$m=EX$。

定义

若$EX^2<\infty$，则X的方差为 $Var(X) := E(X-EX)^2$；标准差为 $\sigma(X) := \sqrt{Var(X)}$。

4、方差性质

注：

- 方差（标准差）反映出随机变量取值的“波动”大小。
- X的方差是X的k阶中心矩$E(X-EX)^k$的特例（$k=2$）。
- X的变异系数为$C_v(X)=\sigma(X)/EX$。

性质

1. $Var(X)=EX^2-(EX)^2\geq 0$;
2. $Var(X)=0\Leftrightarrow P(X=EX)=1$;
3. 若$a,b\in\mathbb{R}$，则$\mathrm{Var}(aX+b)=a^2Var(X)$。

证明：（1）

$$\begin{aligned}Var(X)&=E(X-EX)^2=E[X^2-2XEX+(EX)^2]\\&=EX^2-2EX\cdot EX+(EX)^2=EX^2-(EX)^2\text{。}\end{aligned}$$

5、Chebyshev不等式

定理

设随机变量X满足$EX^2 < +\infty$。则对任意$\varepsilon > 0$，

$$P(|X - EX| \geq \varepsilon) \leq Var(X)/\varepsilon^2.$$

证明：（仅看连续型，离散型类似）

$$\begin{aligned} P(|X - EX| \geq \varepsilon) &= \int_{|x-EX|\geq\varepsilon} p(x)\mathrm{d}x \\ &\leq \int_{|x-EX|\geq\varepsilon} \frac{(x-EX)^2}{\varepsilon^2} p(x)\mathrm{d}x \\ &\leq \varepsilon^{-2}\int_{\mathbb{R}} (x-EX)^2 p(x)\mathrm{d}x = \frac{Var(X)}{\varepsilon^2}。\end{aligned}$$

6、矩的性质（回顾）

性质

$$EX^2 = 0 \Leftrightarrow P(X=0)=1.$$

证明：（只需证“$\Rightarrow$”）

- 注意：$EX^2 = 0 \Rightarrow EX = 0 \Rightarrow Var(X) = 0$;
- $\{\omega : |X(\omega)| > 0\} = \bigcup_{n=1}^{\infty} \{\omega : |X(\omega)| > 1/n\}$;
- $P(|X| > 0) = P\left(\bigcup_{n=1}^{\infty} \{|X| > 1/n\}\right) \leq \sum_{n=1}^{\infty} P(|X| > 1/n)$

  $\leq \sum_{n=1}^{\infty} n^2 Var(X) = 0$。（Chebyshev不等式）
- 所以$P(X=0)=1$。

7、二点分布

二点分布$X \sim b(1,p)$：

X	0	1
P	$1-p$	p

，$X^2 \sim b(1,p)$。

所以

- $EX = p = P(X=1)$，
- $Var(X) = EX^2 - (EX)^2 = p - p^2 = p(1-p)$。

注：

- 事件A的概率$P(A)$等于相应的示性函数的数学期望EI_A。
- 这说明人们可摈弃Kolmogorov的概率公理化定义，先建立随机变量数学期望的公理系统，然后再导出随机事件的概率！

8、二项分布$b(n,p)$

X	0	1	$\dots$	k	$\dots$	n
P	$(1-p)^n$	$np(1-p)^{n-1}$	$\dots$	$\binom{n}{k}p^k(1-p)^{n-k}$	$\dots$	p^n

$$
\begin{aligned}
EX &= \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k}\\
&= \sum_{k=0}^{n} k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}\\
&= np\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}(1-p)^{n-k}\\
&= np\sum_{i=0}^{n-1} \frac{(n-1)!}{i!(n-1-i)!}p^{i}(1-p)^{n-1-i} \qquad （令 i=k-1）\\
&= np[p+(1-p)]^{n-1} = np。
\end{aligned}
$$

9、二项分布$b(n,p)$

$$
\begin{aligned}
Var(X) &= EX^2-(EX)^2 = E[X(X-1)]+EX-(EX)^2\\
E[X(X-1)] &= \sum_{k=0}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k}\\
&= n(n-1)p^2\sum_{k=2}^{n}\binom{n-2}{k-2}p^{k-2}(1-p)^{n-k}\\
&= n(n-1)p^2,\\
Var(X) &= E[X(X-1)]+EX-(EX)^2\\
&= n(n-1)p^2+np-(np)^2\\
&= np(1-p).
\end{aligned}
$$

10、Poisson分布

定义

若$\lambda > 0$且$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$，$k=0,1,2,\dots$，则称$X$是服从参数为$\lambda$的*Poisson*分布，简记为$X \sim P(\lambda)$。

$$
\begin{aligned}
EX &= \sum_{k=0}^{\infty} kP(X=k) = \sum_{k=1}^{\infty} k\lambda^k e^{-\lambda}/k! \\
&= \lambda e^{-\lambda}\sum_{k=1}^{\infty} \lambda^{k-1}/(k-1)! = \lambda; \\
Var(X) &= E[X(X-1)] + EX - (EX)^2 \\
&= \sum_{k=2}^{\infty} k(k-1)\lambda^k e^{-\lambda}/k! + \lambda - \lambda^2 \\
&= \lambda^2 e^{-\lambda}\sum_{k=2}^{\infty} \lambda^{k-2}/(k-2)! + \lambda - \lambda^2 = \lambda.
\end{aligned}
$$

11、Poisson定理

定理

考虑二项分布$b(n, p_n)$，设$\lim\limits_{n\to\infty} np_n = \lambda > 0$，

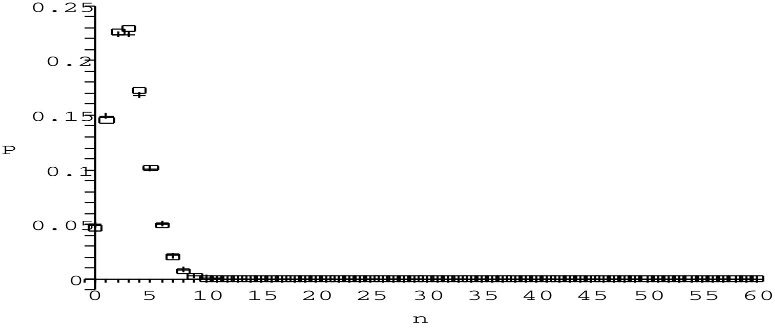
则$p_k = \binom{n}{k} p_n^k (1-p_n)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}, \qquad n \to \infty$。

证明：

$$\begin{aligned} p_k &= \frac{n(n-1)\cdots(n-k+1)}{k!n^k}(np_n)^k\left(1-\frac{np_n}{n}\right)^{n-k} \\ &= \left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\frac{[\lambda+o(1)]^k}{k!}\left[1-\frac{\lambda+o(1)}{n}\right]^{n-k} \\ &\to \frac{\lambda^k}{k!}e^{-\lambda}, \qquad n\to\infty. \end{aligned}$$

12、Poisson定理

- Poisson定理结果表明：当n很大、p很小、np近似为正常数λ时，二项分布$b(n,p)$中的概率值可用参数为λ的Poisson分布的相应概率值近似。
- 下图中的方格表示$b(60,0.05)$的概率分布，十字表示$\lambda = 60 \times 0.05 = 3$的Poisson分布的概率分布。



13、Poisson定理应用

例 保险公司接受理赔案件的次数

区间$[0,\infty)$表示时间，对任何区间$A\subset[0,\infty)$，记$N(A)$表示在时段A内发生的理赔案件的次数。假设

1. 在互不相交的时间段$A_1,\ldots,A_n$中,理赔的发生是相互独立的;
2. 在任意时间段$(t,t+h]$中，

$$P(N(t,t+h]=1)=\lambda h+o(h),\qquad P(N(t,t+h]\geq 2)=o(h),$$

即在很短的一段时间内，有理赔案件发生的概率近似和时间段长度h成正比（比例系数λ不依赖时刻t的值），于是，在这很短的一段时间内出现多起理赔案件的可能性不大。

求保险理赔次数$N(s,s+t]$的统计规律。

14、Poisson定理应用

考虑保险理赔次数，$P(N(s,s+t]=k)$。

- n等分时间段$(s,s+t]$，得到n个长度为$\frac{t}{n}$的小时间段。
- 在时间段$(s,s+t]$内刚好有k件理赔发生的情形分为两类：
  1. 刚好有k个小时间段使得每个时间段内恰有一次理赔；
  2. 至少有一个小时间段中发生了至少两次理赔。
- 记这两类事件发生的概率分别为P_1, P_2。
- 则$P(N(s,s+t]=k)=P_1+P_2$，其中

$$P_1=\binom{n}{k}\left[\lambda\frac{t}{n}+o\left(\frac{t}{n}\right)\right]^k\left[1-\lambda\frac{t}{n}+o\left(\frac{t}{n}\right)\right]^{n-k},$$

$$P_2=o\left(\frac{t}{n}\right)_1+o\left(\frac{t}{n}\right)_2+\cdots+o\left(\frac{t}{n}\right)_n=o(1)。$$

- 让$n\to\infty$，由Poisson定理得

$$P(N(s,s+t]=k)=\frac{(\lambda t)^k}{k!}e^{-\lambda t}.$$

- 当然可以类似求可靠性的方法建立微分方程求解.你试试看?

15、几何分布 $G(p)$

X	1	2	$\cdots$	n	$\cdots$
P	p	pq	$\cdots$	pq^{n-1}	$\cdots$

$$
\begin{aligned}
EX &= \sum_{n=1}^{\infty} npq^{n-1} = \sum_{n=1}^{\infty}\left(\sum_{k=1}^{n} pq^{n-1}\right) \\
&= \sum_{k=1}^{\infty}\left(\sum_{n=k}^{\infty} pq^{n-1}\right) \qquad (\ldots\text{求和交换次序!}) \\
&= \sum_{k=1}^{\infty} q^{k-1} = \frac{1}{p}.
\end{aligned}
$$

解释： 如果我们掷一次骰子得1点的可能性为1/6，那么我们可以期望平均6次左右独立连续抛掷中应该能掷得一次1点。

思考： 若掷硬币得正面概率p，则平均掷几次可得“正反”？

16、几何分布

$$
\begin{aligned}
E[X(X-1)] &= \sum_{n=1}^{\infty} n(n-1)pq^{n-1} = \sum_{n=1}^{\infty}\left(2\sum_{k=1}^{n-1} kpq^{n-1}\right) \\
&= 2\sum_{k=1}^{\infty} k\left(\sum_{n=k+1}^{\infty} pq^{n-1}\right) = 2\sum_{k=1}^{\infty} kq^{k} \\
&= 2q/p\sum_{k=1}^{\infty} kpq^{k-1} = 2q/pEX = 2q/p^2, \\
Var(X) &= E[X(X-1)] + EX - (EX)^2 \\
&= 2q/p^2 + 1/p - 1/p^2 = q/p^2.
\end{aligned}
$$

17、几何分布的无记忆性

称离散随机变量X具有无记忆性，若：

$$P(X > m+n \mid X > m) = P(X > n), \quad \forall m, n \in \mathbb{N}。 \qquad （1）$$

- 直观含义：在已知前m次试验均未成功的条件下，一直到第$m+n$次也未成功相当于在从第m次试验后重新开始的新过程中前n次均未成功，而此前m次的经历可被完全忘掉。
- （1）$\Leftrightarrow P(X > m+n) = P(X > m)P(X > n)$ （2）
- 而对$X \sim G(p)$，$P(X > n) = q^n$。由此易见等式(2)成立。

反之，若X只取自然数值，且具有无记忆性，则$X \sim G(p)$。

- 事实上，（2）$\Leftrightarrow P(X > n) = q^n$，其中$q = P(X > 1)$，
- 于是，$P(X = n) = P(X > n-1) - P(X > n) = pq^{n-1}$ （$p = 1 - q$）。

18、超几何分布

定义

设有N件产品，其中有M件次品。若从中不放回地随机抽取n件，则其中含有的次品数X服从超几何分布，记为$X \sim h(n, N, M)$。

(注：若为放回抽样，则$X \sim b(n, M/N)$。)

$$
\begin{aligned}
P(X=k) &= \binom{M}{k}\binom{N-M}{n-k}\Big/\binom{N}{n}, \\
k &\in \mathbb{N} \cap [a,b], a = \max\{0, n-N+M\}, b = \min\{M, n\}; \\
EX &= \sum_{k=a}^{b} k\binom{M}{k}\binom{N-M}{n-k}\Big/\binom{N}{n} \\
&= \frac{Mn}{N}\sum_{k=a}^{b}\binom{M-1}{k-1}\binom{N-M}{n-k}\Big/\binom{N-1}{n-1} \\
&= Mn/N \qquad \text{（习题1.2，1.（5））}
\end{aligned}
$$

19、超几何分布

$$
\begin{aligned}
P(X=k) &= \binom{M}{k}\binom{N-M}{n-k}\Big/\binom{N}{n}, \quad k\in\mathbb{N}\cap[a,b], \\
E[X(X-1)] &= \sum_{k=a}^{b} k(k-1)\binom{M}{k}\binom{N-M}{n-k}\Big/\binom{N}{n} \\
&= M(M-1)\sum_{k=a}^{b}\binom{M-2}{k-2}\binom{N-M}{n-k}\Big/\binom{N}{n} \\
&= M(M-1)\binom{N-2}{n-2}\Big/\binom{N}{n} \\
&= \frac{M(M-1)n(n-1)}{N(N-1)};
\end{aligned}
$$

20、超几何分布

$$
\begin{aligned}
EX &= Mn/N, \\
E[X(X-1)] &= \frac{M(M-1)n(n-1)}{N(N-1)}, \\
Var(X) &= E[X(X-1)] + EX - (EX)^2 \\
&= \frac{M(M-1)n(n-1)}{N(N-1)} + \frac{Mn}{N} - \left(\frac{Mn}{N}\right)^2 \\
&= \frac{nM(N-M)(N-n)}{N^2(N-1)}.
\end{aligned}
$$

注：当$n \ll N$时， $h(n,N,M) \approx b(n,M/N)$.