数学期望引例1

- 衡量某地区的收入情况——
  可用平均收入作为指标；
- 对班级或个人学习情况进行比较——
  可用平均成绩作为指标...
- 某射手进行实弹射击，每次射击的命中环数是随机变量X，现在他射击了100次，结果如下，如何评价他的射击水平?

命中环数x_i	10	9	8	7	6	5
频数μ_i	20	30	15	10	15	10

数学期望引例1

如何评价他的射击水平?

命中环数x_i	10	9	8	7	6	5
频数μ_i	20	30	15	10	15	10

计算他各次射击结果的算术平均值:

$$\frac{1}{100}\sum_{i=1}^{6} x_i\mu_i = \sum_{i=1}^{6} x_i \frac{\mu_i}{100} \Rightarrow = \sum_{i=1}^{6} x_i f_i \Rightarrow = \sum_{i=1}^{n} x_i p_i$$

$$= 10\times 0.2 + 9\times 0.3 + 8\times 0.15 + 7\times 0.1 + 6\times 0.15 + 5\times 0.1$$

$$= 8$$

可作为随机变量的数字特征!

数学期望引例2

设随机变量X的分布律如下，计算其均值

$$P\left\{X=(-1)^k\frac{2^k}{k}\right\}=\frac{1}{2^k}\quad(k=1,2,\cdots)$$

$$\sum_{k=1}^{\infty}x_k p_k=\sum_{k=1}^{\infty}(-1)^k\frac{2^k}{k}\cdot\frac{1}{2^k}=\sum_{k=1}^{\infty}(-1)^k\frac{1}{k}=-\ln 2$$

思考： 先计算正数部分，再计算负数部分，然后合并，结果如何？

分开计算，级数不收敛——改变求和的顺序，导致结果不一致！

数学期望引例2

设随机变量X的分布律如下，计算其均值

$$P\left\{X=(-1)^k\frac{2^k}{k}\right\}=\frac{1}{2^k}\quad(k=1,2,\cdots)$$

分开计算，级数不收敛——改变求和的顺序，导致结果不一致！

作为随机变量的数字特征应是唯一的，不该因求和顺序不同而不一致！

为避免求和顺序不同导致的结果不一致，应增加限制条件

期望是否存在？

1) 设$R.V.X$服从拉普拉斯分布,其概率密度为:

$$f(x)=\frac{1}{2}e^{-|x|} \qquad -\infty < x < +\infty \quad 求\ E(X)$$

2) 设$R.V.X$服从柯西分布,其概率密度为:

$$f(x)=\frac{1}{\pi\left(1+x^{2}\right)} \qquad -\infty < x < +\infty \quad 求\ E(X)$$

解: 1) $E(X)=\frac{1}{2}\int_{-\infty}^{+\infty} x e^{-|x|}dx=0$

被积函数为奇函数且积分区间关于原点对称

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

在一般的$E(X)$计算中,为简便记,我们常省略事先判断绝对收敛这一步.

泊松分布期望值

1. $X \sim P(\lambda)$, 则 $E(X)=\lambda$

证明: $P\{X=k\}=e^{-\lambda}\dfrac{\lambda^k}{k!}\quad k=0,1,2,\ldots.$

$$E(X)=\sum_{k=0}^{+\infty}k\cdot P\{X=k\}=\sum_{k=0}^{+\infty}ke^{-\lambda}\frac{\lambda^k}{k!}$$

$$=\sum_{k=1}^{+\infty}\frac{\lambda^k}{(k-1)!}e^{-\lambda}\ \overset{\text{令}m=k-1}{=\!=\!=\!=}\ \sum_{m=0}^{+\infty}\frac{\lambda^{m+1}}{m!}e^{-\lambda}$$

$$=\lambda\sum_{m=0}^{+\infty}\frac{\lambda^m}{m!}e^{-\lambda}$$

$$=\lambda$$

注意到它与泊松分布相似，但泊松分布取值是从0开始！

要点：

1、按照定义写出期望的表达式；

2、将被加部分化成常见分布的分布律形式；

3、利用分布律之和等于1化简。

2. $X \sim B(n,p)$, 则 $E(X) = np$

证明: $P\{X=i\} = C_n^i (1-p)^{n-i} p^i \quad i = 0,1,2,\ldots, n$

$$E(X) = \sum_{i=0}^{n} i C_n^i (1-p)^{n-i} p^i$$

$$C_n^i = \frac{n!}{i!(n-i)!}$$

$$= \sum_{i=1}^{n} n C_{n-1}^{i-1} (1-p)^{n-1-(i-1)} p^i$$

$$\underline{\underline{令 k = i-1}}\ np \sum_{k=0}^{n-1} C_{n-1}^{k} (1-p)^{n-1-k} p^k$$

$$= np[p + (1-p)]^{n-1}$$

$$= np$$

正态分布期望值

3. $X \sim N(\mu, \sigma^2)$, 则 $E(X) = \mu$

证明: $E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \dfrac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$

$$\underline{\underline{t = \frac{x-\mu}{\sigma}}} \quad \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} (\mu + \sigma t) e^{-\frac{t^2}{2}} dt$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \sigma t e^{-\frac{t^2}{2}} dt + \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \mu e^{-\frac{t^2}{2}} dt$$

$$= 0 + \mu \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= \mu$$

随机变量函数的数学期望

一个庄家在一个签袋中放有8个白、8个黑的围棋子。规定：每个摸彩者交一角钱作“手续费”，然后一个从袋中摸出五个棋子，按下面“摸子中彩表”给“彩金”。

摸到	五个白	四个白	三个白	其它
彩金	2元	2角	5分	共乐一次

随机变量函数的数学期望

设X表示摸出的白围棋子个数，Y表示一次得到的彩金，则X，Y的分布律为

X	0	1	2	3	4	5
P	0.0128	0.1282	0.3590	0.3590	0.1282	0.0128
Y	0	0	0	0.05	0.2	2
P	0.0128	0.1282	0.3590	0.3589	0.1282	0.0128

随机变量函数的数学期望

X	0	1	2	3	4	5
P	0.0128	0.1282	0.3590	0.3590	0.1282	0.0128
Y	0	0	0	0.05	0.2	2
P	0.0128	0.1282	0.3590	0.3589	0.1282	0.0128

设$Y=g(X)$，则

$$E(Y)=g(0)\times 0.0128+g(1)\times 1282+g(2)\times 0.3590$$

$$+g(3)\times 0.3590+g(4)\times 0.1282+g(5)\times 0.0128$$

$$=0\times 0.5001$$

$$+0.05\times 0.3589+0.2\times 0.1282\quad +2\times 0.0128$$

$$=\sum_{i=1}^{6} g(x_i)\cdot P\{X=x_i\}$$

设随机变量X的数学期望存在，证明:

$$E\left\{[X-E(X)]^2\right\}=E(X^2)-[E(X)]^2$$

设随机变量X的概率密度为

$$f(x)=\begin{cases}1+x & -1\le x<0\\ 1-x & 0\le x<1\\ 0 & \text{其它}\end{cases}$$

试求 $E\left\{[X-E(X)]^2\right\}$

证明: $E\left\{[X-E(X)]^2\right\}=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)dx$

$$=\int_{-\infty}^{+\infty}\left\{x^2-2xE(X)+[E(X)]^2\right\}f(x)dx$$

$$=E(X^2)-[E(X)]^2$$

$\because E(X)=0$ （参见书上 例4.1.3）

$$E\left\{\left[X-E(X)\right]^2\right\}=E\left(X^2\right)-\left[E(X)\right]^2$$

$$=E\left(X^2\right)$$

$$=\int_{-\infty}^{+\infty}x^2f(x)dx$$

$$=\int_{-1}^{0}x^2(1+x)dx+\int_{0}^{1}x^2(1-x)dx$$

$$=1/6$$

$X\cdot Y$的期望

设X, Y相互独立， $E(X), E(Y)$存在, 且

$$P\{X = x_i\} = p_{i.}\ , i = 1,2,\cdots$$
$$P\{Y = y_j\} = p_{.j}\ , j = 1,2,\cdots$$

求 $E(XY)$

解: $$E(XY) = \sum_i \sum_j x_i y_j P\{X = x_i, Y = y_j\}$$
$$= \sum_i \sum_j x_i y_j P_{i.} P_{.j}$$
$$= \sum_i x_i P_{i.} \sum_j y_j P_{.j}$$
$$= E(X)E(Y)$$

$\min(X,Y)$ 的 期 望

X,Y相互独立,且服从$N(0,1)$分布. 试求$E[\min(X,Y)]$

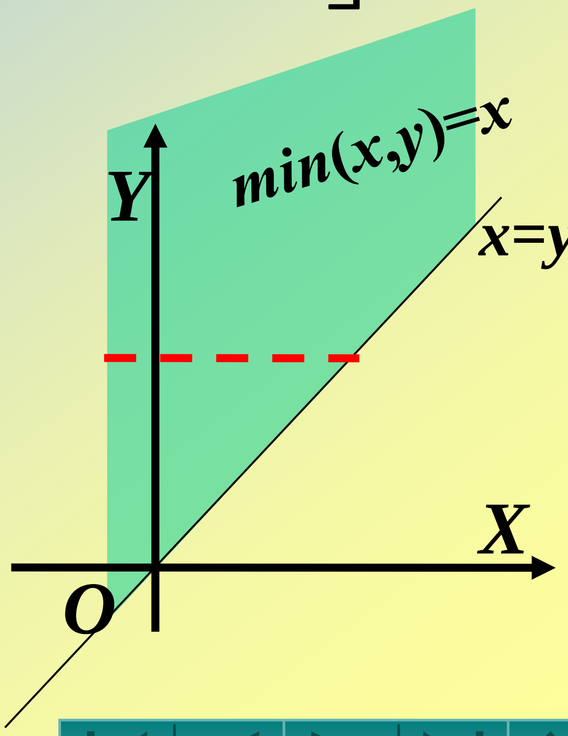
解: $E[\min(X,Y)] = \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}\min(x,y)f(x,y)dx\right]dy$

$$= \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{y} xf(x,y)dx\right]dy + \int_{-\infty}^{+\infty}\left[\int_{y}^{+\infty} yf(x,y)dx\right]dy$$

又 $\because f(x,y)$关于x,y对称

$$\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{x} yf(x,y)dy\right]dx$$

$$= \int_{-\infty}^{+\infty}\left[\int_{y}^{+\infty} yf(x,y)dx\right]dy$$



X, Y相互独立,且服从$N(0,1)$分布. 试求$E[\min(X,Y)]$

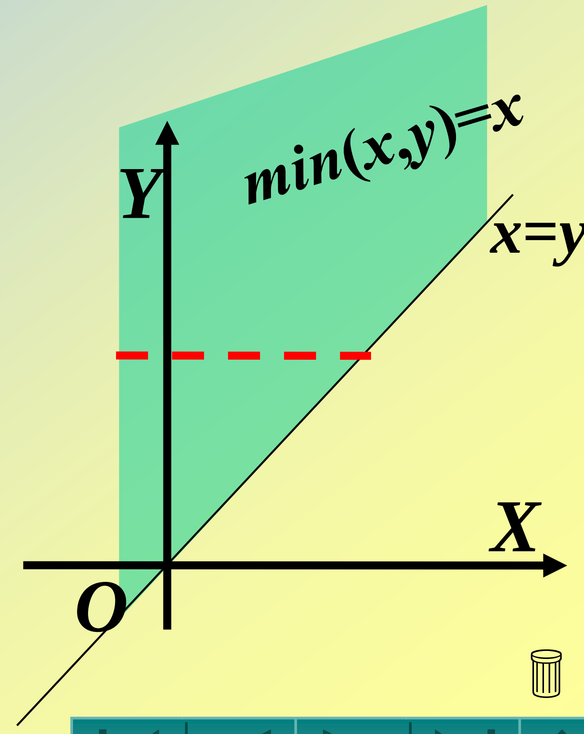
解: $$E[\min(X,Y)] = \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}\min(x,y)f(x,y)dx\right]dy$$

$$= 2\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{y} xf(x,y)dx\right]dy$$

$$= 2\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{y} x\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dx\right]dy$$

$$= \int_{-\infty}^{\infty} -\frac{e^{-y^2}}{\pi}dy$$

$$= -\frac{1}{\sqrt{\pi}}$$



练习解答

设随机变量X与Y相互独立,且$X,Y\sim N\left(0,\frac{1}{2}\right)$

则 $E\left(|X-Y|\right)=$ ________

解法一: $f(x,y)=f_X(x)f_Y(y)=\dfrac{1}{\pi}e^{-\left(x^2+y^2\right)}\qquad (x,y)\in R^2$

$$E\left(|X-Y|\right)=\iint_{R^2}|x-y|f(x,y)d\sigma=\cdots\cdots\cdots=\sqrt{\frac{2}{\pi}}$$

解法二:利用正态分布可加性

$Y\sim N\left(0,{}^{1}/_{2}\right)\quad\Rightarrow\quad -Y\sim N\left(0,\dfrac{1}{2}\right)$ 【见例题3.4.7结论】

$X\sim N\left(0,\dfrac{1}{2}\right)$

$\because$ X、Y相互独立，由正态分布可加性

【见例题3.4.11结论】

$$\Rightarrow X-Y=X+(-Y)\sim N(0,1)$$

令$Z=X-Y\sim N(0,1)$，则

$$E(|X-Y|)=E(|Z|)=\int_{-\infty}^{+\infty}|z|\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz$$

$$=\frac{1}{\sqrt{2\pi}}\cdot 2\int_0^{+\infty}ze^{-\frac{z^2}{2}}dz=\sqrt{\frac{2}{\pi}}\int_0^{+\infty}e^{-\frac{z^2}{2}}d\,\frac{z^2}{2}$$

$$=\sqrt{\frac{2}{\pi}}\left[-e^{-\frac{z^2}{2}}\right]_0^{+\infty}=\sqrt{\frac{2}{\pi}}$$

2） $E(cX+b) = cE(X)+b$

(仅就X为连续型的情况给出证明)

$$E(cX+b)=\int_{-\infty}^{+\infty}(cx+b)f(x)dx$$

$$=\int_{-\infty}^{+\infty}cxf(x)dx+\int_{-\infty}^{+\infty}bf(x)dx$$

$$=c\int_{-\infty}^{+\infty}xf(x)dx+b\int_{-\infty}^{+\infty}f(x)dx$$

$$=cE(X)+b$$

证明:

$$E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)$$

证明:左边

$$=E\{XY-YE(X)-XE(Y)+E(X)E(Y)\}$$

$$=E(XY)-E(X)E(Y)-E(X)E(Y)+E(X)E(Y)$$

$$=E(XY)-E(X)E(Y)$$

调整设备的平均次数

某批产品的次品率为0.1,检验员每天检查4次,每次随机地取10件产品进行检验,如果发现其中的次品数多于1 ,就去调整设备. 以X表示一天中调整设备的次数,试求 $E(X)$

解: 设Y_i为第i次检查时发现的次品数,则

$$Y_i \sim B(10, 0.1)$$

设X_i为第i次检查时需要调整设备的次数

$$X_i = \begin{cases} 1 & Y_i > 1 \\ 0 & Y_i \le 1 \end{cases}$$ 显然有 $$X = \sum_{i=1}^{4} X_i$$

解: 设Y_i为第i次检查时发现的次品数,则

$$Y_i \sim B(10, 0.1)$$

设X_i为第i次检查时需要调整设备的次数

$$X_i = \begin{cases} 1 & Y_i > 1 \\ 0 & Y_i \le 1 \end{cases}$$ 显然有 $$X = \sum_{i=1}^{4} X_i$$

$$P\{X_i = 0\} = P\{Y_i = 0\} + P\{Y_i = 1\}$$

$$= (0.9)^{10} + C_{10}^{1}(0.1)(0.9)^9 = 1.9(0.9)^9$$

$$P\{X_i = 1\} = P\{Y_i > 1\} = 1 - P\{Y_i \le 1\}$$

$$= 1 - 1.9(0.9)^9$$

则 X_i服从0-1分布

X_i	0	1
p	$1.9(0.9)^9$	$1-1.9(0.9)^9$

$$E(X_i)=1-1.9(0.9)^9 \quad i=1,2,3,4$$

从而 $E(X)=\sum_{i=1}^{4}E(X_i)=4\left[1-1.9(0.9)^9\right]=1.0556$

例:超几何分布的数学期望

求服从超几何分布的随机变量X的期望$E(X)$

$$P\{X=m\}=C_M^m C_{N-M}^{n-m}/C_N^n \quad m=0,1,2,\ldots,n \quad n\le M\le N$$

原始模型： N个球中有M个红球，余下为白球，从中任取n个球，n个球中的红球数为X

分析

1）显然直接求解很困难。因此应该想到用数学期望的性质求解。

求服从**超几何分布**的随机变量X的期望$E(X)$

$$P\{X=m\}=C_M^m C_{N-M}^{n-m}/C_N^n \quad m=0,1,2,\dots,n \quad n\le M\le N$$

原始模型： N个球中有M个红球，余下为白球，从中任取n个球，n个球中的红球数为X

分析

2）可以设想这n个球是逐个不放回抽取的

令X_i表示第i次取到红球的个数，$i=1,2,\cdots,n$

则 $X=X_1+X_2+\cdots+X_n$

3) 由抽签的公平性有：

$$P\{X_i=1\}=M/N$$

解：设想这n个球是逐个不放回抽取的，共取了n次。令X_i表示第i次取到红球的个数，$i = 1, 2, \cdots, n$　则

$$X = X_1 + X_2 + \cdots + X_n$$

由抽签的公平性有：$P\{X_i = 1\} = M/N$

从而$E(X_i) = 1 \times \frac{M}{N} + 0 \times \left(1 - \frac{M}{N}\right) = \frac{M}{N}$

$$E(X) = \sum_{i=1}^{n} E(X_i) = \frac{nM}{N}$$

M, N较大时，超几何分布近似为二项分布：

$$X \sim B\left(n, \frac{M}{N}\right)$$

方法总结：

1. 设置X_i，使得$X = X_1 + X_2 + \cdots + X_n$

2. 分析X_i的所有可能取值(一般只有两个取值)

3. 按照X_i的分布律计算出$E(X_i)$

4. $E(X) = E(X_1) + \cdots + E(X_n)$

例:射击命中次数的期望

向某一目标进行射击，直至命中k次为止。已知命中率为$p>0$. 求命中次数X的数学期望。

分析：X的分布律为

$$P\{X=i\}=C_{i-1}^{k-1}p^k(1-p)^{i-k},\qquad i=k,k+1,\cdots$$

$$E(X)=\sum_{i=k}^{+\infty} iC_{i-1}^{k-1}p^k(1-p)^{i-k}$$

直接计算是一件很困难的事。因此考虑用数学期望的性质$X=X_1+X_2+\cdots+X_n$进行求解。

$1\quad 2\quad .\quad .\quad i\quad .\quad .\quad k$

X_i表示第$i-1$次命中以后，到第i次命中的射击次数

$$X=X_1+X_2+\cdots+X_n$$

X_i的分布律为:

X_i	1	2	$\cdots$	m	$\cdots$
$P\{X_i = m\}$	p	$(1-p)p$	$\cdots$	$(1-p)^{m-1}p$	$\cdots$

$$E(X_i) = \sum_{m=1}^{+\infty} m(1-p)^{m-1} p$$

$$= p\left[\sum_{m=1}^{+\infty} m x^{m-1}\right]_{x=1-p} = p\left[\sum_{m=1}^{+\infty} x^{m}\right]'_{x=1-p}$$

例:4.1.7 向某一目标进行射击，直至命中k次为止。已知命中率为$p>0$. 求命中次数X的数学期望。

解：设X_i表示第$i-1$次命中以后，到第i次命中的射击次数。则有$X=X_1+X_2+\cdots+X_n$

X_i的分布律为：

X_i	1	2	…..	m	……
$P\{X_i=m\}$	p	$(1-p)p$	…..	$(1-p)^{m-1}p$	……

$$E(X_i)=\sum_{m=1}^{+\infty}m(1-p)^{m-1}p=p\left[\sum_{m=1}^{+\infty}mx^{m-1}\right]_{x=1-p}=p\left[\sum_{m=1}^{+\infty}x^{m}\right]'_{x=1-p}=\frac{1}{p}$$

从而 $E(X)=E(X_1)+E(X_2)+\ldots+E(X_k)$

$$=\frac{1}{p}k$$

传染病检查

若某地流行甲肝，为控制病情，急需对人群作血清检查。若共有n个人，人群的发病率为p，有两种检查方式：①一人一次，共需n次　② k个人一组，若血清为阴性，则查一次即可，否则需对每人重作一次检查，此时共需 $k+1$次。

问：选用哪一种方法更好?

•[分析] 用第一种方法需n 次检查，现判断用第二种方法需多少次。

•第二种方法所用次数显然是不定的，设为X次，则我们可据其期望$E(X)$ 来判断这种方法是否会更好。

X的期望直接计算比较困难，我们设第i个人需作X_i次检查$(i=1,\cdots,n)$，则$X=X_1+\cdots+X_n$

$$E(X)=E(X_1)+\cdots+E(X_n)$$

- X_i 的可能取值有两个：$1/k$ 或 $1+1/k$ (k 个人共作$k+1$次，每人平均$(1+k)/k$次)

- 令A_j表示k个人中第j个人患病，($j=1,\cdots,k$) 则

$$P\left\{X_i=\frac{1}{k}\right\}=P\left(\overline{A_1}\cap\overline{A_2}\cap\cdots\cap\overline{A_k}\right)=P\left(\overline{A_1}\right)\cdot P\left(\overline{A_2}\right)\cdot\cdots\cdot P\left(\overline{A_k}\right)$$

$$=(1-p)^k$$

$$P\left\{X_i=1+\frac{1}{k}\right\}=P\left(A_1\cup A_2\cdots\cup A_k\right)=1-(1-p)^k$$

X_i的分布律为：

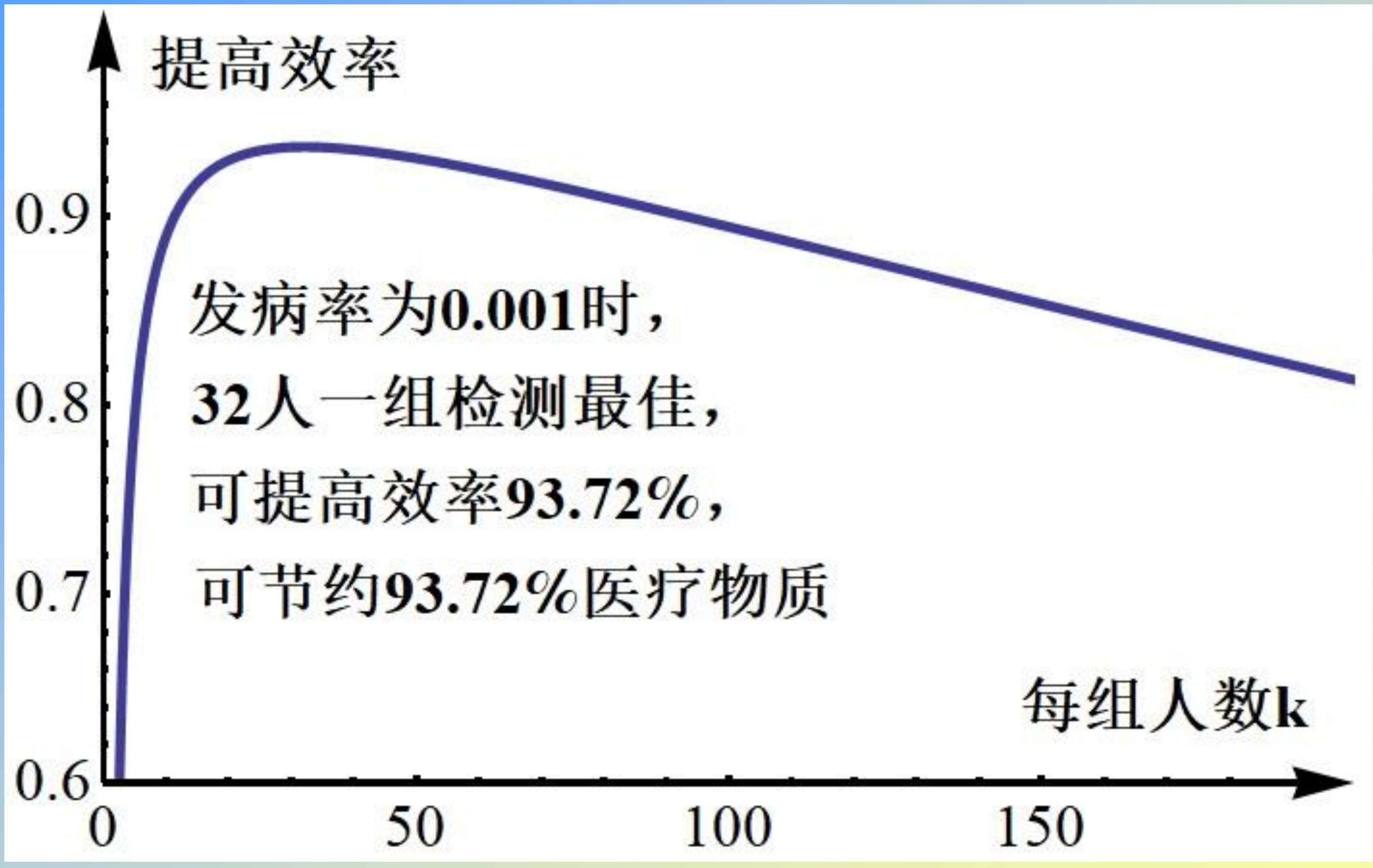
X_i	$\frac{1}{k}$	$1+\frac{1}{k}$
P	$(1-p)^k$	$1-(1-p)^k$

$$E(X_i)=\frac{1}{k}(1-p)^k+\left(1+\frac{1}{k}\right)\left[1-(1-p)^k\right]$$

$$=1+\left[\frac{1}{k}-(1-p)^k\right]=1-\boxed{\left[(1-p)^k-\frac{1}{k}\right]}$$

$E(X)=n-n\times\left[(1-p)^k-\frac{1}{k}\right]$，故关键在于 $(1-p)^k-\frac{1}{k}$ 的取值，当它大于0时，可减少检查次数!

例如，$p=0.01$时，取$k=10$，则 $(1-p)^k-\frac{1}{k}=0.8044$，可减少百分之八十的工作量!



提高效率
0.9
0.8
0.7
0.6
发病率为0.001时，
32人一组检测最佳，
可提高效率93.72%，
可节约93.72%医疗物质
每组人数k
0
50
100
150