Probability and Statistics - P59

负二项随机变量(negative binomial random variable)

考虑多次独立重复的伯努利试验，每次试验成功率为 $0<p<1$

对于正整数 r，令随机变量 X 表示持续试验直到累计 r 次成功时的总试验次数

则对于其概率分布列 $p(n)=\Pr\{X=n\}$，$n=r, r+1, r+2, \cdots$

如果第 n 次试验中取得第 r 次试验成功

则前 $n-1$ 次试验中包含且仅包含 $r-1$ 次试验成功

于是有 $\Pr\{X=n\}=\binom{n-1}{r-1}p^{r-1}(1-p)^{n-r}\cdot p=\binom{n-1}{r-1}p^r(1-p)^{n-r}$

$$\sum_{n=r}^{\infty}\Pr\{X=n\}=\sum_{n=r}^{\infty}\binom{n-1}{r-1}p^{r}(1-p)^{n-r}$$
$$=p^r\sum_{n=r}^{\infty}\binom{(n-r)+r-1}{n-r}(1-p)^{n-r}$$
$$=p^r\sum_{k=0}^{\infty}\binom{k+r-1}{k}(1-p)^k=p^r\sum_{k=0}^{\infty}\underbrace{\binom{-r}{k}}_{\text{广义二项式系数}}(-1)^k(1-p)^k$$
$$=p^r\cdot\frac{1}{[1-(1-p)]^r}$$
$$=1$$

称随机变量 X 服从参数为 (r,p) 的负二项分布

特别地当 $r=1$ 时，随机变量 X 退化为参数为 p 的几何分布

考虑独立重复的成功率为 p 的伯努利试验，第 r 次成功出现在第 m 次失败之前的概率

于是可知 第 r 次成功出现在第 m 次失败之前 当且仅当

第 r 次成功出现在第 $r+m-1$ 次试验或之前的试验

则有 $\Pr\{$第 r 次成功出现在第 m 次失败之前$\}=\sum_{n=r}^{r+m-1}\Pr\{X=n\}$

$$=\sum_{n=r}^{r+m-1}\binom{n-1}{r-1}p^r(1-p)^{n-r}$$

对于服从参数为 (r,p) 的负二项分布的随机变量 X

有 X 的 k 阶矩 $E[X^k]=\sum_{n=r}^{\infty}n^k\binom{n-1}{r-1}p^r(1-p)^{n-r}$

$$=\sum_{n=r}^{\infty}n^k\cdot\frac{(n-1)!}{(r-1)!(n-r)!}p^r(1-p)^{n-r}$$
$$=\frac{r}{p}\sum_{n=r}^{\infty}n^{k-1}\frac{n!}{r!(n-r)!}p^{r+1}(1-p)^{n-r}$$
$$=\frac{r}{p}\sum_{m=r+1}^{\infty}(m-1)^{k-1}\binom{m-1}{r}p^{r+1}(1-p)^{m-(r+1)}$$
$$=\frac{r}{p}E[(Y-1)^{k-1}]$$

其中随机变量 Y 服从参数为 $(r+1,p)$ 的负二项分布

当 $k=1$ 时，随机变量 X 的期望为 $E[X]=\frac{r}{p}E[(Y-1)^0]=\frac{r}{p}$

当 $k=2$ 时，$E[X^2]=\frac{r}{p}E[(Y-1)^1]=\frac{r}{p}(E[Y]-1)$

$$=\frac{r}{p}\left(\frac{r+1}{p}-1\right)$$

则随机变量 X 的方差 $Var[X]=E[X^2]-(E[X])^2=\frac{r}{p}\left(\frac{r+1}{p}-1\right)-\left(\frac{r}{p}\right)^2$

$$=\frac{r(1-p)}{p^2}$$

Probability and Statistics - P60

巴拿赫火柴问题(Banach's match box problem)，考虑数学家的左右口袋各有一盒火柴

每次随机地从一个口袋的火柴盒中取出一根使用

假设开始时两盒火柴各有N根，则当数学家发现其中一盒为空时

考虑另一盒火柴恰有k根剩余的概率，其中$k=1,2,\cdots,N$

令事件E_k表示当数学家发现左盒为空时，右盒中剩余k根火柴

则E_k发生当且仅当第$N+1+N-k$次选中左边的口袋

且前$N+N-k$次选择中有N次选中左边口袋

即 $\Pr\{E_k\} = \binom{N+N-k}{N}\left(\frac{1}{2}\right)^N\left(\frac{1}{2}\right)^{N-k}\cdot\frac{1}{2}$

$= \binom{2N-k}{N}\cdot\left(\frac{1}{2}\right)^{2N-k+1}$

又令事件F_k表示当数学家发现右边为空时，左边剩余k根火柴

可知 $\Pr\{F_k\} = \Pr\{E_k\} = \binom{2N-k}{N}\cdot\left(\frac{1}{2}\right)^{2N-k+1}$

取随机变量X表示当数学家发现其中一盒为空时，另一盒中剩余的火柴数量

又对于任意 $k=1,2,\cdots,N$，E_k与F_k是互不相容的

于是 $\Pr\{X=k\} = \Pr\{E_k\}+\Pr\{F_k\} = \binom{2N-k}{N}\left(\frac{1}{2}\right)^{2N-k}$

考虑开始时其中一盒包含N_1根火柴，另一盒包含N_2根火柴，并平凡地假定$N_1 \le N_2$

令事件E_k表示发现第一盒为空时，第二盒中剩余k根火柴

事件F_k表示发现第二盒为空时，第一盒中剩余k根火柴

则当$k=1,2,\cdots,N_1$时

$\Pr\{E_k\} = \binom{N_1+N_2-k}{N_1}\left(\frac{1}{2}\right)^{N_1+N_2-k+1}$

$\Pr\{F_k\} = \binom{N_1+N_2-k}{N_2}\left(\frac{1}{2}\right)^{N_1+N_2-k+1}$

而当$k=N_1+1,\cdots,N_2$时，$\Pr\{F_k\}=0=\binom{N_1+N_2-k}{N_2}\left(\frac{1}{2}\right)^{N_1+N_2-k+1}$

$\Pr\{E_k\} = \binom{N_1+N_2-k}{N_1}\left(\frac{1}{2}\right)^{N_1+N_2-k+1}$

于是有 $\Pr\{X=k\} = \Pr\{E_k\}+\Pr\{F_k\}$

$=\left[\binom{N_1+N_2-k}{N_1}+\binom{N_1+N_2-k}{N_2}\right]\left(\frac{1}{2}\right)^{N_1+N_2-k+1}$

考虑当其中一盒为空时，另一盒包含k根火柴，其中$k=1,2,\cdots,N$

事件E_k发生当且仅当第$N+N-k$次选中为空的火柴盒

且前$N+N-k-1$次选择中有$N-1$次选中为空的火柴盒

则 $\Pr\{X=k\} = \binom{2}{1}\cdot\binom{N+N-k-1}{N-1}\cdot\left(\frac{1}{2}\right)^{N-1}\left(\frac{1}{2}\right)^{N-k}\cdot\frac{1}{2}$

$=\binom{2N-k-1}{N-1}\cdot\left(\frac{1}{2}\right)^{2N-k-1}$

Probability and Statistics - P61

对于服从参数为 (r, p) 的负二项分布的随机变量 X

和服从参数为 (n, p) 的二项分布的随机变量 Y，其中 $0<p<1$

则有 $Pr\{X>n\} = Pr\{Y<r\}$

证明过程有，当 $n<r$ 时，$Pr\{X>n\} = 1 = Pr\{Y<r\}$

当 $n \geq r$ 时，则考虑 $Pr\{X>n\}$

可以描述为在前 n 次试验中只有不超过 $r-1$ 次成功

于是有 $Pr\{X>n\} = \sum_{i=0}^{r-1} \binom{n}{i} p^i (1-p)^{n-i}$

$= Pr\{Y<r\}$

又 $Pr\{X>n\} = \sum_{i=n+1}^{\infty} \binom{i-1}{r-1} p^r (1-p)^{i-r}$

于是有 $\sum_{i=n+1}^{\infty} \binom{i-1}{r-1} p^r (1-p)^{i-r} = \sum_{i=0}^{r-1} \binom{n}{i} p^i (1-p)^{n-i}$

对于服从参数为 (n, p) 的二项随机变量 $B(n,p)$

则有 $Pr\{B(n,p) \leq i\} = 1 - Pr\{B(n,1-p) \leq n-i-1\}$，其中 $i=0,1,\cdots,n$

证明过程有，首先有 $1 - Pr\{B(n,1-p) \leq n-i-1\}$

$= Pr\{B(n,1-p) > n-i-1\} = Pr\{B(n,1-p) \geq n-i\}$

而 $Pr\{B(n,p) \leq i\}$ 可描述为成功次数不大于 i 的概率

又等价于失败次数不小于 $n-i$ 的概率

而失败次数服从参数为 $(n, 1-p)$ 的二项分布

于是有 $Pr\{B(n,p) \leq i\} = Pr\{B(n,1-p) \geq n-i\}$

$= 1 - Pr\{B(n,1-p) \leq n-i-1\}$

超几何随机变量 (hyper geometric random variable)

考虑一个坛子中共有 N 个球，其中包含 m 个白球和 $N-m$ 个黑球，其中 $0 \leq m \leq N, N \in N^+$

从坛子中随机无放回地取出 n 个球，$n \in N^+$，且有 $0<n \leq N$

令随机变量 X 表示取出来的白球数量

则考虑 $Pr\{X=i\}$，其中 $i=0,1,\cdots,n$

从 m 个白球中取出 i 个的不同方式数为 $\binom{m}{i}$

从 $N-m$ 个黑球中取出 $n-i$ 个的不同方式数为 $\binom{N-m}{n-i}$

而从 N 个球中选择 n 个的不同方式数为 $\binom{N}{n}$

于是有 $Pr\{X=i\} = \binom{m}{i}\binom{N-m}{n-i} / \binom{N}{n}$，其中 $i=0,1,\cdots,n$

特别地有仅当 $\max\{0, n+m-N\} \leq i \leq \min\{n, m\}$ 时

有 $Pr\{X=i\} > 0$

称随机变量 X 服从参数为 (N, m, n) 的超几何分布，记为 $X \sim Hypergeometric(N, m, n)$

Probability and Statistics - P62

超几何随机变量，考虑对于N个球，其中有m个白球和$N-m$个黑球，其中m,N为正整数，且有$m\le N$

从中无放回地随机抽取n个球，考虑其中白球的数量

令随机变量X表示其中白球的数量，可知服从参数(N,m,n)的超几何分布

于是对于$i=0,1,2,\cdots,n$，

$$\Pr\{X=i\}=\binom{m}{i}\binom{N-m}{n-i}\Big/\binom{N}{n}$$
$$=\frac{m!}{i!(m-i)!}\times\frac{(N-m)!}{(n-i)!(N-m-n+i)!}\Big/\frac{N!}{n!(N-n)!}$$
$$=\frac{n!}{i!(n-i)!}\times\frac{m!}{(m-i)!}\times\frac{(N-m)!}{(N-m-n+i)!}\Big/\frac{N!}{(N-n)!}$$
$$=\binom{n}{i}\times[m(m-1)\cdots(m-i+1)]\times[(N-m)(N-m-1)\cdots(N-m-n+i+1)]$$
$$/[N(N-1)\cdots(N-n+1)]$$
$$=\binom{n}{i}\times\frac{m(m-1)\cdots(m-i+1)}{N(N-1)\cdots(N-i+1)}\times\frac{(N-m)(N-m-1)\cdots[N-m-(n-i)+1]}{(N-i)(N-i-1)\cdots[N-i-(n-i)+1]}$$

注意到当$N,m\gg n,i$时，即有$\frac{n,i}{N,m}\to 0$

则有
$$\frac{m(m-1)\cdots(m-i+1)}{N(N-1)\cdots(N-i+1)}\approx\left(\frac{m}{N}\right)^i$$
$$\frac{(N-m)(N-m-1)\cdots[N-m-(n-i)+1]}{(N-i)(N-i-1)\cdots[N-i-(n-i)+1]}\approx\left(\frac{N-m}{N}\right)^{n-i}$$

[redacted]再令$p=\frac{m}{N}$，则有$1-p=\frac{N-m}{N}$

于是有$\Pr\{X=i\}\approx\binom{n}{i}p^i(1-p)^{n-i}$，即近似地服从参数为$(n,p)$的二项分布

[redacted] 可以描述为，当坛子中的球数量N与白球数量m均远大于取出的球数量n时

则可以认为之前取的球不影响其后取球的概率

考虑在某个区域内[redacted]统计某个物种的个体总数N，对于未知的N进行极大似然估计

首先捕捉m只物种个体并进行标记然后放回，并假设一段时间后这些个体充分地分散

然后再次捕捉n只物种个体，令随机变量X表示捕捉到的标记的个体数

假设两次捕捉时物种的总数N没有变化，且捕捉到每只个体[redacted]都是等概率的

可知随机变量X服从参数为(N,m,n)的超几何分布

再令$P_i(N)$表示该区域实际上包含N个个体的条件下获得观测值$X=i$的概率

则对于给定的观测值$X=i$，使$P_i(N)$达到最大值的N值应是一个合理估计

于是有$P_i(N)\equiv\Pr\{X=i\}=\binom{m}{i}\binom{N-m}{n-i}\Big/\binom{N}{n}$

取
$$P_i(N)/P_i(N-1)=\left[\binom{m}{i}\binom{N-m}{n-i}\Big/\binom{N}{n}\right]\Big/\left[\binom{m}{i}\binom{N-1-m}{n-i}\Big/\binom{N-1}{n}\right]$$
$$=\left[\binom{N-m}{n-i}\Big/\binom{N-1-m}{n-i}\right]\Big/\left[\binom{N}{n}\Big/\binom{N-1}{n}\right]$$
$$=\left[\frac{(N-m)!}{(n-i)!(N-m-n+i)!}\Big/\frac{(N-m-1)!}{(n-i)!(N-m-n+i-1)!}\right]\Big/\left[\frac{N!}{n!(N-n)!}\Big/\frac{(N-1)!}{n!(N-n-1)!}\right]$$
$$=\frac{N-m}{N-m-n+i}\Big/\frac{N}{N-n}$$
$$=\frac{(N-m)(N-n)}{N(N-m-n+i)}$$

当$P_i(N)/P_i(N-1)\ge 1$时，$\frac{(N-m)(N-n)}{N(N-m-n+i)}\ge 1$

则有$N\le mn/i$，即$P_i(N)$先[redacted]递增再递减

且[redacted]在$N=$[redacted]$\lfloor mn/i\rfloor$处，$P_i(N)$取得最大值

Probability and Statistics - P63

超几何随机变量，对于服从参数为 (N,m,n) 的超几何分布的随机变量 X

有概率分布列 $p(i) = Pr\{X=i\} = \binom{m}{i}\binom{N-m}{n-i}/\binom{N}{n}$，$i=0,1,\cdots,n$

则 $\sum_{i=0}^{n} p(i) = \sum_{i=0}^{n}\binom{m}{i}\binom{N-m}{n-i}/\binom{N}{n}$

$= [\sum_{i=0}^{n}\binom{m}{i}\binom{N-m}{n-i}]/\binom{N}{n}$

根据范德蒙德恒等式 (Vandermonde's identity)

有 $\sum_{i=0}^{n}\binom{m}{i}\binom{N-m}{n-i} = \binom{m+(N-m)}{i+(n-i)} = \binom{N}{n}$

于是有 $\sum_{i=0}^{n} p(i) = [\sum_{i=0}^{n}\binom{m}{i}\binom{N-m}{n-i}]/\binom{N}{n} = \binom{N}{n}/\binom{N}{n} = 1$

对于服从参数为 (N,m,n) 的超几何分布的随机变量 X

有 X 的 k 阶矩 $E[X^k] = \sum_{i=0}^{n} i^k Pr\{X=i\}$

$= \sum_{i=1}^{n} i^k\binom{m}{i}\binom{N-m}{n-i}/\binom{N}{n}$

$= \sum_{i=1}^{n} n[i\binom{m}{i}]i^{k-1}\binom{N-m}{n-i}/[n\binom{N}{n}]$

$= \sum_{i=1}^{n} n[m\binom{m-1}{i-1}]i^{k-1}\binom{N-m}{n-i}/[N\binom{N-1}{n-1}]$

$= \frac{nm}{N}\sum_{i=1}^{n} i^{k-1}\binom{m-1}{i-1}\binom{N-m}{n-i}/\binom{N-1}{n-1}$

$= \frac{nm}{N}\sum_{i=0}^{n-1}(i+1)^{k-1}\binom{m-1}{i}\binom{(N-1)-(m-1)}{n-1-i}/\binom{N-1}{n-1}$

$= \frac{nm}{N}E[(Y+1)^{k-1}]$

其中随机变量 Y 服从参数为 $(N-1, m-1, n-1)$ 的超几何分布

当 $k=1$ 时，有 X 的期望 $E[X] = \frac{nm}{N}E[(Y+1)^0] = \frac{nm}{N}$

当 $k=2$ 时，有 X 的二阶矩 $E[X^2] = \frac{nm}{N}E[(Y+1)^1] = \frac{nm}{N}E[Y+1]$

$= \frac{nm}{N}[\frac{(n-1)(m-1)}{N-1}+1]$

则有 X 的方差为 $Var(X) = E[X^2]-(E[X])^2$

$= \frac{nm}{N}[\frac{(n-1)(m-1)}{N-1}+1-\frac{nm}{N}]$

再令 $p=\frac{m}{N}$，即有 $0\le p\le 1$

则有 X 的期望 $E[X] = \frac{nm}{N} = np$

X 的方差 $Var(X) = \frac{nm}{N}[\frac{(n-1)(m-1)}{N-1}+1-\frac{nm}{N}]$

$= np[(n-1)\frac{Np-1}{N-1}+1-np]$

$= np[(n-1)\frac{Np-p+(p-1)}{N-1}+1-np]$

$= np[(n-1)(p+\frac{p-1}{N-1})+1-np]$

$= np[1-p-(n-1)\frac{1-p}{N-1}]$

$= np(1-p)\frac{N-n}{N-1}$

于是可知服从参数 (N,m,n) 的超几何分布期望等于服从参数 $(n,\frac{m}{N})$ 的二项分布期望

而当 $N \gg n$ 时，有 $\frac{n}{N}\to 0$，有 $\frac{N-n}{N-1}\to 1$

于是有 $Var(X)\approx np(1-p)$，即约等于服从参数 $(n,\frac{m}{N})$ 的二项分布方差

Probability and Statistics - P64

超几何随机变量，对于服从参数为(N, m, n)的超几何分布的随机变量X，考虑其单调性

对于 $Pr\{X=k+1\}/Pr\{X=k\} = \left[\binom{m}{k+1}\binom{N-m}{n-k-1}/\binom{N}{n}\right]/\left[\binom{m}{k}\binom{N-m}{n-k}/\binom{N}{n}\right]$

$= \left[\frac{m!}{(k+1)!(m-k-1)!}\cdot\frac{(N-m)!}{(n-k-1)!(N-m-n+k+1)!}\right]/\left[\frac{m!}{k!(m-k)!}\cdot\frac{(N-m)!}{(n-k)!(N-m-n+k)!}\right]$

$= \frac{k!}{(k+1)!}\cdot\frac{(m-k)!}{(m-k-1)!}\cdot\frac{(n-k)!}{(n-k-1)!}\cdot\frac{(N-m-n+k)!}{(N-m-n+k+1)!}$

$= (m-k)(n-k)/(k+1)(N-m-n+k+1)$

则当 $Pr\{X=k+1\}/Pr\{X=k\} \geq 1$ 时

$(m-k)(n-k)/(k+1)(N-m-n+k+1) \geq 1$

有 $k \leq (mn+m+n-N-1)/(N+2)$

于是可知随机变量X的概率分布列首先单调递增，再单调递减

并在 $\lfloor(mn+m+n-N-1)/(N+2)\rfloor+1 = \lfloor\frac{(m+1)(n+1)}{N+2}\rfloor$ 处取得最大值

对于包含N个白球和M个黑球的坛子，每次从中无放回地取出一个球，其中 $N\in N^+$, $M\in N^+$

令$p(m,n)$表示在取出m个黑球之前已经取出n个白球的概率，其中 $1\leq m\leq M$, $n\leq N$

对于给定的正整数m，令随机变量X_m表示取出m个黑球前已经取出的白球个数

则可知 $p(m,n) = Pr\{X_m=n\}$，其中 $n=0,1,\cdots,N$

又 $X_m=n$ 当且仅当前 $m+n-1$ 次取球中取得了 $m-1$个黑球与n个白球

且第$m+n$次取球取出的是黑球

可知前$m+n-1$次取球中取得的白球数服从参数为$(N+M, N, m+n-1)$的超几何分布

于是有取出n个白球的概率为 $\binom{N}{n}\binom{M}{m-1}/\binom{N+M}{n+m-1}$

而第$m+n$次取球取得黑球的概率为 $(M-m+1)/(N+M-n-m+1)$

于是有 $p(m,n)=Pr\{X_m=n\} = \left[\binom{N}{n}\binom{M}{m-1}/\binom{N+M}{n+m-1}\right]\cdot\left[(M-m+1)/(N+M-n-m+1)\right]$

$= \binom{N}{n}\left[(M-m+1)\binom{M}{M-m+1}\right]/\left[(N+M-n-m+1)\binom{N+M}{N+M-n-m+1}\right]$

$= \binom{N}{n}\left[M\binom{M-1}{M-m}\right]/\left[(N+M)\binom{N+M-1}{N+M-n-m}\right]$

$= \frac{M}{N+M}\binom{N}{n}\binom{M-1}{m-1}/\binom{N+M-1}{n+m-1}$

注意到对于 $p(m,n) = \frac{M}{N+M}\binom{N}{n}\binom{M-1}{m-1}/\binom{N+M-1}{n+m-1}$

可以拆分为两部分的乘积，$\frac{M}{N+M}$ 与 $\binom{N}{n}\binom{M-1}{m-1}/\binom{N+M-1}{n+m-1}$

则存在另一种描述方式，对于在取出m个黑球之前已经取出n个白球

即首先从全部$N+M$个球中取出一个黑球并视为第$n+m$次取球的结果

则其概率为 $\binom{M}{1}/\binom{N+M}{1} = \frac{M}{N+M}$

则其余部分为从$M-1$个黑球与N个白球中取出$n+m-1$个球的问题

则取出的白球数服从参数为$(N+M-1, N, n+m-1)$的超几何分布

于是取出n个白球的概率为 $\binom{N}{n}\binom{M-1}{m-1}/\binom{N+M-1}{n+m-1}$

于是有 $p(m,n) = \frac{M}{N+M}\binom{N}{n}\binom{M-1}{m-1}/\binom{N+M-1}{n+m-1}$

Probability and Statistics - P65

负超几何随机变量(negative hypergeometric distribution)

考虑在一个坛子中包含n个白球和m个黑球，其中 $n\in N^+$，$m\in N$

从中不放回地随机取球，直到取出k个白球，其中 $k\le n$

令随机变量X表示总共取球的次数，考虑 $Pr\{X=r\}$

注意和负二项分布的区别在于各次取球的概率不同

可知对于 $X=r$，其中 $k\le r\le k+m$，

当且仅当在前 $r-1$ 次取球中包含 $k-1$ 个白球和 $r-k$ 个黑球

且第r次取球取到白球

又从n个白球和m个黑球中取 $r-1$ 个球中白球数量

服从参数为 $(n+m, n, r-1)$ 的超几何分布

则取得 $k-1$ 个白球的概率为 $\binom{n}{k-1}\binom{m}{r-k}/\binom{n+m}{r-1}$

而在第r次取球前，在 $n+m-r+1$ 个球中包含 $n-k+1$ 个白球

于是有 $Pr\{X=r\} = \frac{n-k+1}{n+m-r+1}\cdot\binom{n}{k-1}\binom{m}{r-k}/\binom{n+m}{r-1}$

$$= [(n-k+1)\binom{n}{n-k+1}]\binom{m}{r-k}/[(n+m-r+1)\binom{n+m}{n+m-r+1}]$$

$$= [n\binom{n-1}{n-k}]\binom{m}{r-k}/[(n+m)\binom{n+m-1}{n+m-r}]$$

$$= \frac{n}{n+m}\binom{n-1}{k-1}\binom{m}{r-k}/\binom{n+m-1}{r-1}$$

于是可以描述为，首先从 $n+m$ 个球中取出一个白球，并视为第r次取球

再从其余的 $n-1$ 个白球与m个黑球中取出 $r-1$ 个球，其中包含 $k-1$ 个白球

于是服从参数为 $(n+m-1, n-1, r-1)$ 的超几何分布

则有 $Pr\{X=r\} = \frac{n}{n+m}\binom{n-1}{k-1}\binom{m}{r-k}/\binom{n+m-1}{r-1}$

再考虑概率分布列的其他形式有

$$Pr\{X=r\} = \frac{n}{n+m}\binom{n-1}{k-1}\binom{m}{r-k}/\binom{n+m-1}{r-1}$$

$$= \frac{n}{n+m}\times\frac{(n-1)!}{(k-1)!(n-k)!}\times\frac{m!}{(r-k)!(m-r+k)!}\Big/\frac{(n+m-1)!}{(r-1)!(n+m-r)!}$$

$$= \frac{(r-1)!}{(k-1)!(r-k)!}\times\frac{(n+m-r)!}{(n-k)!(m-r+k)!}\Big/\frac{(n+m)!}{n!\,m!}$$

$$= \binom{r-1}{k-1}\binom{n+m-r}{n-k}/\binom{n+m}{n}$$

可以描述为，对n个白球与m个黑球进行排列，并考虑第r个位置是第k个白球

则前 $r-1$ 个位置中有 $k-1$ 个白球，后 $n+m-r$ 个位置中有 $n-k$ 个白球

于是有 $Pr\{X=r\} = \binom{r-1}{k-1}\binom{n+m-r}{n-k}/\binom{n+m}{n}$

由于对于 $\binom{n+m}{n}$ 中的任意一种排列第k个白球的位置 $r=k, k+1, \cdots, k+m$

于是有 $\sum_{r=k}^{k+m} Pr\{X=r\} = \sum_{r=k}^{k+m}\binom{r-1}{k-1}\binom{n+m-r}{n-k}/\binom{n+m}{n} = 1$

于是称随机变量X服从参数为 $(n+m, n, k)$ 的负超几何随机变量

记为 $X\sim H^-(n+m, n, k)$

Probability and Statistics - P66

负超几何随机变量，对于服从参数为$(n+m, n, k)$的负超几何随机分布的随机变量X

有概率分布率 $\Pr\{X=r\} = \binom{r-1}{k-1}\binom{n+m-r}{n-k}/\binom{n+m}{n}$，其中 $r=k, k+1, \cdots, k+m$

注意到由于有 $\sum_{r=k}^{k+m}\binom{r-1}{k-1}\binom{n+m-r}{n-k}/\binom{n+m}{n} = 1$

于是可以得到组合恒等式，对于 $n\in N^+$，$m\in N$，正整数 $k\leq n$

有 $\sum_{r=k}^{k+m}\binom{r-1}{k-1}\binom{n+m-r}{n-k} = \binom{n+m}{n}$

考虑随机变量X的原点升阶s阶矩，其中 $s\in N^+$，

有 $X^{(s)} = X(X+1)\cdots(X+s-1) = X\cdot(X+1)^{(s-1)}$，特别地令 $X^{(0)}=1$

则有
$$E[X^{(s)}] = \sum_{r=k}^{k+m} r^{(s)}\binom{r-1}{k-1}\binom{n+m-r}{n-k}/\binom{n+m}{n}$$
$$= \sum_{r=k}^{k+m}(r+1)^{(s-1)}\left[r\binom{r-1}{k-1}\right]\binom{n+m-r}{n-k}/\binom{n+m}{n}$$
$$= \sum_{r=k}^{k+m}(r+1)^{(s-1)}\left[k\binom{r}{k}\right]\frac{(n+m-r)!}{(n-k)!(m-r+k)!}\Big/\frac{(n+m)!}{n!\,m!}$$
$$= k\sum_{r=k}^{k+m}(r+1)^{(s-1)}\binom{r}{k}\frac{[(n+1)+m-(r+1)]!}{[(n+1)-(k+1)]!\,[m-(r+1)+(k+1)]!}\Big/\frac{(n+m)!}{n!\,m!}$$

取 $r'=r+1$，$k'=k+1$

则
$$E[X^{(s)}] = k\sum_{r'=k'}^{k'+m} r'^{(s-1)}\binom{r'-1}{k'-1}\frac{[(n+1)+m-r']!}{[(n+1)-k']!\,(m-r'+k')!}\cdot\frac{n+m+1}{n+1}\Big/\left[\frac{(n+m)!}{n!\,m!}\cdot\frac{n+m+1}{n+1}\right]$$
$$= \frac{k(n+m+1)}{n+1}\cdot\sum_{r'=k'}^{k'+m} r'^{(s-1)}\binom{r'-1}{k'-1}\frac{[(n+1)+m-r']!}{[(n+1)-k']!\,(m-r'+k')!}\Big/\frac{[(n+1)+m]!}{(n+1)!\,m!}$$
$$= \frac{k(n+m+1)}{n+1}E[Y^{(s-1)}]$$

其中随机变量Y服从参数为$(n+m+1, n+1, k+1)$的负超几何随机分布

当$s=1$时，有 $X^{(1)}=X$，$Y^{(0)}=1$

于是随机变量X的期望 $E[X] = E[X^{(1)}] = \frac{k(n+m+1)}{n+1}E[Y^{(0)}] = \frac{k(n+m+1)}{n+1}$

考虑通过随机变量X的原点升阶2阶矩计算随机变量X的方差

首先当$s=2$时，有 $X^{(2)} = X(X+1)$

于是有
$$E[X^{(2)}] = E[X(X+1)] = \frac{k(n+m+1)}{n+1}E[Y^{(1)}]$$
$$= k(k+1)(n+m+1)(n+m+2)/(n+1)(n+2)$$

则随机变量X的方差
$$Var[X] = E[X^2]-(E[X])^2 = E[X^2]+E[X]-(E[X])^2-E[X]$$
$$= E[X(X+1)]-(E[X])^2-E[X]$$
$$= \frac{k(k+1)(n+m+1)(n+m+2)}{(n+1)(n+2)} - \left[\frac{k(n+m+1)}{n+1}\right]^2 - \frac{k(n+m+1)}{n+1}$$
$$= \frac{k(n+m+1)}{(n+1)^2(n+2)}\big[(k+1)(n+1)(n+m+2) - (n+2)[k(n+m+1)+(n+1)]\big]$$
$$= \frac{k(n+m+1)(m+mn-km)}{(n+1)^2(n+2)}$$
$$= \frac{km(n+m+1)(n+1-k)}{(n+1)^2(n+2)}$$

当s为任意正整数时，可以递归地得到任意的随机变量X的原点升阶s阶矩

有 $E[X^{(s)}] = k^{(s)}(n+m+1)^{(s)}/(n+1)^{(s)}$

Probability and Statistics - P67

负超几何随机变量：对于服从参数为$(n+m, n, k)$的负超几何分布的随机变量X

有概率分布列 $Pr\{X=r\} = \binom{r-1}{k-1}\binom{n+m-r}{n-k}/\binom{n+m}{n}$，其中 $r=k, k+1, \cdots, k+m$

又
$$Pr\{X=r\} = \binom{r-1}{k-1}\binom{n+m-r}{n-k}/\binom{n+m}{n}$$
$$= \binom{r-1}{k-1}\frac{(n+m-r)!}{(n-k)!(m-r+k)!} \Big/ \frac{(n+m)!}{n!\,m!}$$
$$= \binom{r-1}{k-1}\frac{n!}{(n-k)!}\cdot\frac{m!}{(m-r+k)!} \Big/ \frac{(n+m)!}{(n+m-r)!}$$
$$= \binom{r-1}{k-1}\frac{n(n-1)\cdots(n-k+1)\times m(m-1)\cdots(m-r+k+1)}{(n+m)(n+m-1)\cdots(n+m-r+1)}$$
$$= \binom{r-1}{k-1}\frac{n(n-1)\cdots(n-k+1)}{(n+m)(n+m-1)\cdots(n+m-k+1)}\cdot\frac{m(m-1)\cdots(m-r+k+1)}{(n+m-k)(n+m-k-1)\cdots(n+m-k-r+k+1)}$$

注意当 $n, m >> r, k$ 时，即有 $\frac{r, k}{n, m} \to 0$

则有
$$\frac{n(n-1)\cdots(n-k+1)}{(n+m)(n+m-1)\cdots(n+m-k+1)} \approx \left(\frac{n}{n+m}\right)^k$$
$$\frac{m(m-1)\cdots(m-r+k+1)}{(n+m-k)(n+m-k-1)\cdots(n+m-k-r+k+1)} \approx \left(\frac{m}{n+m}\right)^{r-k}$$

再令 $p=\frac{n}{n+m}$，则有 $(1-p)=\frac{m}{n+m}$

则有 $Pr\{X=r\} \approx \binom{r-1}{k-1}p^k(1-p)^{r-k}$

即随机变量X近似地服从参数为(k, p)的负二项分布

可以描述为当白球数量n与黑球数量m远大于取球数量时

则之前取球的结果不影响之后取球的概率

又有随机变量X的期望 $E[X] = \frac{k(n+m+1)}{n+1}$

方差 $Var(X) = \frac{km(n+m+1)(n+1-k)}{(n+1)^2(n+2)}$

则当 $n, m >> r, k$ 时，即 $\frac{r, k}{n, m} \to 0$，而 $p=\frac{n}{n+m}$，$(1-p)=\frac{m}{n+m}$

有 $E[X] = \frac{k(n+m+1)}{n+1} \approx \frac{k}{p}$

$$Var(X) = \frac{km(n+m+1)(n+1-k)}{(n+1)^2(n+2)} = k\cdot\frac{m}{n+m}\cdot\frac{n+m}{n+1}\cdot\frac{n+m+1}{n+1}\cdot\frac{n+1-k}{n+2} \approx \frac{k(1-p)}{p^2}$$

即此时随机变量X近似地服从参数为$(k, \frac{n}{n+m})$的负二项分布

假设坛子中包含n个球，编号为$1, 2, \cdots, n$，从中有放回地随机从坛子中取m次球

令随机变量X表示抽取的m次球中出现的最大编号，

考虑随机变量X的概率分布列 $Pr\{X=k\}$，其中 $k=1, 2, \cdots, n$

首先有 $Pr\{X=k\} = Pr\{X\leq k\} - Pr\{X<k\}$
$$= Pr\{X\leq k\} - Pr\{X\leq k-1\}$$

又 $Pr\{X\leq k\}$ 可以描述为m次取球均为编号为$1, 2, \cdots, k$中的球

又取球是有放回的，所以每次都取到$1, 2, \cdots, k$的概率是相同

即有 $Pr\{X\leq k\} = \left(\frac{k}{n}\right)^m$

于是有 $Pr\{X=k\} = Pr\{X\leq k\} - Pr\{X\leq k-1\}$
$$= \left(\frac{k}{n}\right)^m - \left(\frac{k-1}{n}\right)^m，其中\ m\in N^+，k=1, 2, \cdots, n$$

Probability and Statistics - P68

假设坛子中有N个球，编号为$1, 2, \cdots, N$，其中 $N \in N^+$

从中无放回地随机取出n个球，其中 $1 \le n \le N$

令随机变量Y表示取出的n个球中最大的编号，则考虑随机变量Y的概率分布列

对于 $Pr\{Y=k\}$，其中 $k=n, n+1, \cdots, N$

注意取出的球中最大的编号为k发生当且仅当

取出了编号为k的球且从编号$1, 2, \cdots, k-1$的球中取出其余$n-1$个球

又从N个球中取出n个球总共有$\binom{N}{n}$种不同方式

于是可知 $Pr\{Y=k\} = \binom{k-1}{n-1} / \binom{N}{n}$，其中 $k=n, n+1, \cdots, N$

则有 $\sum_{k=n}^{N} Pr\{Y=k\} = \sum_{k=n}^{N} \binom{k-1}{n-1} / \binom{N}{n}$

$= \sum_{k=0}^{N-n} \binom{n-1+k}{n-1} / \binom{N}{n}$

$= \binom{n-1+N-n+1}{N-n} / \binom{N}{n}$

$= \binom{N}{N-n} / \binom{N}{n} = 1$

再考虑随机变量Y的期望$E[Y]$

有 $E[Y] = \sum_{k=n}^{N} k \binom{k-1}{n-1} / \binom{N}{n} = \sum_{k=n}^{N} n \binom{k}{n} / \binom{N}{n}$

$= n \sum_{k'=n'}^{N+1} \binom{k'-1}{n'-1} / \binom{N}{n}$，令 $k' = k+1$，$n' = n+1$

$= n \cdot \binom{N+1}{n'} / \binom{N}{n} = n \cdot \binom{N+1}{n+1} / \binom{N}{n}$

$= \dfrac{n(N+1)}{n+1}$

假设有n个芯片，从中有放回地随机抽取，直到抽到一个曾经抽到过的芯片为止

令随机变量X表示此时的抽取次数，考虑其概率分布列

首先对于 $X=1$，由于在第一次之前未抽取任何芯片，于是 $Pr\{X=1\}=0$

对于 $X=n+1$，根据鸽笼原理，由于只有n块不同芯片

则前$n+1$次抽取中必定存在两次抽取抽到同一个芯片

于是对于任意 $k > n+1$，$Pr\{X=k\}=0$

于是考虑 $Pr\{X=k\}$，其中 $k=2, 3, \cdots, n+1$

注意到当 $X=k$ 发生时，当且仅当

前$k-1$次抽取全部为不同的芯片

第k次抽取得到的是前$k-1$次抽取到的芯片之一

由于抽取是有放回的，

于是 $Pr\{X=k\} = \frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-k+2}{n} \times \frac{k-1}{n}$

$= (k-1)n! / n^k (n-k+1)!$，其中 $k=2, 3, \cdots, n+1$

Probability and Statistics - P69

假设有n个芯片，随机地抽取直到抽取到一个曾经抽到过的芯片

特别地有，抽取出的芯片在下一次抽取之后才放回

令随机变量X表示此时抽取的次数，考虑其概率分布列

首先对于 $X=1$，由于在第一次抽取前没有抽取过任何芯片，于是有 $\Pr\{X=1\}=0$

对于 $X=2$，由于第一次抽取的芯片没有放回，于是第二次抽不可能与第一次相同，于是有 $\Pr\{X=2\}=0$

对于 $X=n+1$，由于只有n个芯片，于是根据鸽笼原理

则前 $n+1$ 次抽取中必定存在两次抽取到同一个芯片

于是对于任意 $k>n+1$，有 $\Pr\{X=k\}=0$

于是考虑 $\Pr\{X=k\}$，$k=3,4,\dots,n+1$，可知 $X=k$ 发生当且仅当

前 $k-1$ 次抽取全部为不同的芯片

第k次抽取到前 $k-2$ 次抽取到的芯片之一，注意第 $k-1$ 次抽取的芯片未放回

则有 $\Pr\{X=k\} = \frac{n}{n} \times \frac{n-1}{n-1} \times \cdots \times \frac{n-k+2}{n-1} \times \frac{k-2}{n-1}$

$= (k-2)(n-1)! / (n-1)^{k-1}(n-k+1)!$

假设坛子初始包含1个红球和1个蓝球，随机地抽取直到抽取到蓝球

特别地有，每次抽取之后，将抽取的球与一个相同颜色的球一起放回

令随机变量X表示此时抽取的次数，考虑其概率分布列

对于 $\Pr\{X=k\}$，其中 $k=1,2,3,\dots$，可知 $X=k$ 发生当且仅当

前 $k-1$ 次抽取全部为红球 且第k次抽取是蓝球

则有 $\Pr\{X=k\} = \left(\frac{1}{2} \times \frac{2}{3} \times \cdots \times \frac{k-1}{k}\right) \times \frac{1}{k+1}$

$= \frac{1}{k} \cdot \frac{1}{k+1} = \frac{1}{k} - \frac{1}{k+1}$

再考虑至少取到i次红球的概率，其中 $i=1,2,\dots$

有 $\Pr\{X>i\} = 1 - \Pr\{X \le i\}$

$= 1 - \sum_{k=1}^{i} \Pr\{X=k\}$

$= 1 - \sum_{k=1}^{i} \left(\frac{1}{k} - \frac{1}{k+1}\right)$

$= 1 - \left(1 - \frac{1}{2}\right) - \left(\frac{1}{2} - \frac{1}{3}\right) - \cdots - \left(\frac{1}{i} - \frac{1}{i+1}\right)$

$= \frac{1}{i+1}$

又 $\sum_{k=1}^{\infty} \Pr\{X=k\} = \lim_{n\to+\infty} \sum_{k=1}^{n} \Pr\{X=k\} = \lim_{n\to+\infty} \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right)$

$= \lim_{n\to+\infty} \left(1 - \frac{1}{n+1}\right) = 1$

于是可知最终取得蓝球的概率为1

随机变量X有期望 $E[X] = \sum_{k=1}^{\infty} k \cdot \Pr\{X=k\} = \sum_{k=1}^{\infty} k \cdot \left(\frac{1}{k} \cdot \frac{1}{k+1}\right)$

$= \sum_{k=1}^{\infty} \frac{1}{k+1} = H_\infty - 1$

$= \infty$

Probability and Statistics - P70

墨犬比乌斯函数(Möbius function)，对于正整数 n，考虑n是否为无平方数因数的数(square-free integer)

即对于正整数 $n>1$ 进行素因子分解 有 $n=p_1p_2\cdots p_k$

其中 $p_1,p_2,\cdots,p_k$ 为不相同的素数

或者描述为对于正整数n有素因子分解数 $n=p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k}$

有 $e_1,e_2,\cdots,e_k$ 均不大于或等于2

则有墨犬比乌斯函数，

$$\mu(n)=\begin{cases}1, & n=1\\(-1)^k, & n\text{为无平方数因数的数，且有素因子分解}n=p_1p_2\cdots p_k\\0, & n\text{有大于1的平方数因数}\end{cases}$$

假设对于给定的正整数 N，随机选择正整数 $1\le x\le N$

令事件E表示对于选择的正整数x，$\mu(x)\ne 0$

考虑当N趋向于无穷大时，事件E的发生概率 $\lim_{N\to+\infty}\Pr\{\mu(x)\ne 0\}$

令事件 F_p 表示对于给定的素数 p，选择的正整数 x 满足 $p^2\nmid x$

注意对于给定的正整数N，有 $\lfloor N/p^2\rfloor$ 个正整数使 $p^2\mid x$ 成立

则对于随机选取的x，F_p发生的概率为 $\Pr\{F_p\}=(N-\lfloor N/p^2\rfloor)/N$

而当$N\to+\infty$时，$\lim_{N\to+\infty}\Pr\{F_p\}=(1-\frac{1}{p^2})$

注意到对于不同的素数 p，q，当正整数N趋向于$+\infty$时

事件 F_p 与 F_q 可以视为相互独立的，

且事件E发生当且仅当所有事件 F_p 均发生，对于p为小于N的素数

于是有 $\lim_{N\to+\infty}\Pr\{\mu(x)\ne 0\}=\lim_{N\to+\infty}\prod_{p\text{为小于}N\text{的素数}}(1-\frac{1}{p^2})=\prod_{p\text{为素数}}(1-\frac{1}{p^2})$

$=1/[\prod_{p\text{为素数}}\frac{1}{1-\frac{1}{p^2}}]$

$=1/[\prod_{p\text{为素数}}(1+\frac{1}{p^2}+\frac{1}{p^4}+\cdots)]$

$=1/[\sum_{n=1}^{\infty}\frac{1}{n^2}]$

对于平方数倒数和，考虑欧拉给出的基于无穷级数的不严密计算方法，但结果是正确的

考虑对正弦函数进行泰勒级数展开，

有 $\sin x=\sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^{2k-1}}{(2k-1)!}=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots$

则 $0=\frac{\sin x}{x}=1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\cdots$ 有解为 $x=n\pi$，其中 $n=\pm1,\pm2,\pm3,\cdots$

于是 $\frac{\sin x}{x}$ 也可以表示为基于解的线性因子的乘积，又常数项为1

则有 $\frac{\sin x}{x}=(1-\frac{x}{\pi})(1+\frac{x}{\pi})(1-\frac{x}{2\pi})(1+\frac{x}{2\pi})(1-\frac{x}{3\pi})(1+\frac{x}{3\pi})\cdots$

$=(1-\frac{x^2}{\pi^2})(1-\frac{x^2}{4\pi^2})(1-\frac{x^2}{9\pi^2})\cdots$

比较等式中x^2的系数，则有 $-\frac{1}{3!}=\sum_{n=1}^{\infty}(-\frac{1}{n^2\pi^2})$

于是有 $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$，即有 $\lim_{N\to+\infty}\Pr\{\mu(x)\ne 0\}=\frac{6}{\pi^2}$

于是可以描述为，在所有正整数中，无平方数因数的正整数的比例为 $\frac{6}{\pi^2}$

Probability and Statistics – P71

黎曼ζ函数 (Riemann zeta function)，对于实部(real part)大于1的复数 s，即 $Re(s)>1$

有 $\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}$

当 $s=1$ 时，黎曼ζ函数退化为调和级数

即 $\zeta(1)=$ [redacted] $\sum_{n=1}^{\infty}\frac{1}{n}=\lim_{N\to+\infty}H_N=\infty$

当 $s=2$ 时，黎曼ζ函数即正整数的平方数倒数和

即 $\zeta(2)=\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$

黎曼ζ函数的欧拉乘积公式 (Euler product formula)

即 $\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_{素数p}\frac{1}{1-p^{-s}}$

证明过程是有，$\zeta(s)=\frac{1}{1^s}+\frac{1}{2^s}+\frac{1}{3^s}+\frac{1}{4^s}+\cdots$

$\frac{1}{2^s}\zeta(s)=\frac{1}{2^s}+\frac{1}{4^s}+\frac{1}{6^s}+\frac{1}{8^s}+\cdots$

于是有 $(1-\frac{1}{2^s})\zeta(s)=\frac{1}{1^s}+\frac{1}{3^s}+\frac{1}{5^s}+\frac{1}{7^s}+\cdots=\sum_{2\nmid n,\,n\in N^+}\frac{1}{n^s}$

$\frac{1}{3^s}(1-\frac{1}{2^s})\zeta(s)=\frac{1}{3^s}+\frac{1}{9^s}+\frac{1}{15^s}+\frac{1}{21^s}+\cdots$

$(1-\frac{1}{3^s})(1-\frac{1}{2^s})\zeta(s)=\frac{1}{1^s}+\frac{1}{5^s}+\frac{1}{7^s}+\frac{1}{11^s}+\cdots=\sum_{2\nmid n,\,3\nmid n,\,n\in N^+}\frac{1}{n^s}$

对于所有素数 p 重复以上步骤，则有

$\prod_{素数p}(1-\frac{1}{p^s})\cdot\zeta(s)=\sum_{n\in N^+,\,\forall 素数p\ p\nmid n}\frac{1}{n^s}=1$

于是有 $\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_{素数p}\frac{1}{1-p^{-s}}$

齐夫定律 (Zipf's law)，也称ζ分布或Zipf分布，对于 N 个结果中的某一个发生的概率

[redacted] 有正整数 N : number of element

正整数 $1\le k\le N$: 发生概率从高到低排序的 rank

参数 s : value of exponent charactering distribution，且 $s\geq 0$

则令随机变量 X 表示 [redacted] rank 为 k 的结果出现

同时 $Pr\{X=k\}$ 可以描述为 normalized frequency of element of rank k out of population of N elements

于是有 $Pr\{X=k\}=f(k;s,N)=\frac{1}{k^s}/(\sum_{n=1}^{N}\frac{1}{n^s})$

称随机变量 X 服从参数为 ([redacted] s,N) 的Zip分布

注意到当随机变量 X 的样本空间为正整数时，即有 $N\to+\infty$，另外有 $s>1$

有 $Pr\{X=k\}=\frac{1}{k^s}/(\sum_{n=1}^{\infty}\frac{1}{n^s})=\frac{1}{k^s\zeta(s)}$ 其中 $\zeta(s)$ 为黎曼ζ函数

注意对于Zipf分布，始终有众数(mode)为1，即始终有 $Pr\{X=1\}$ 最大

Probability and Statistics - P72

对于有限或可数无限的样本空间S，可能出现的试验结果 $s \in S$

给定随机变量X，令 $X(s)$ 表示试验结果为 $s \in S$ 时随机变量X的取值

令 $p(s) = \Pr\{s\}$ 表示s作为随机试验结果的概率

则任意事件 $A \subseteq S$ 为有限个或可数无限个互不相容事件 $\{s\}$ 的并

于是有 $\Pr\{A\} = \sum_{s \in A} p(s)$

则有 $E[X] = \sum_{s \in S} X(s)p(s)$

证明过程有，对于随机变量X的不同可能取值 x_i，$i = 1, 2, 3, \cdots$

则对于任意 $i \in N^+$，令 S_i 表示使 $X(s) = x_i$ 的结果集合

即 $S_i = \{s \in S \mid X(s) = x_i\}$

于是有 $E[X] = \sum_i x_i \cdot \Pr\{X = x_i\} = \sum_i x_i \cdot \Pr\{S_i\}$

$= \sum_i x_i \left(\sum_{s \in S_i} p(s)\right)$

且 $\bigcup_i S_i = S$ $\quad = \sum_i \sum_{s \in S_i} x_i \cdot p(s)$

其中 $S_1, S_2, \cdots$ 为互不相容事件 $\quad = \sum_i \sum_{s \in S_i} X(s)p(s)$ [redacted]

[redacted] $\quad = \sum_{s \in S} X(s)p(s)$

可以描述为，$E[X]$ 表示随机变量X的可能取值的加权平均

其中X的每个可能取值的权重为其取得的试验结果的概率

即 $E[X]$ 应该等于 $X(s)$ 的加权平均，其中 $s \in S$

$X(s)$ 的权重为s作为试验[redacted]结果的概率

对于有限或可数无限的样本空间S，可能出现的试验结果 $s \in S$

给定n个随机变量 $X_1, X_2, \cdots, X_n$

令 $X_i(s)$ 表示试验结果为 $s \in S$ 时随机变量 X_i 的取值，其中 $i = 1, 2, \cdots, n$

则有 $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$

证明过程有，令随机变量 $Z = \sum_{i=1}^{n} X_i$

于是 $Z(s)$ 表示试验结果为 $s \in S$ 时随机变量Z的[redacted]取值

则 $E\left[\sum_{i=1}^{n} X_i\right] = E[Z] = \sum_{s \in S} Z(s)p(s)$

$= \sum_{s \in S} [X_1(s) + X_2(s) + \cdots + X_n(s)]\, p(s)$

$= \sum_{s \in S} [X_1(s)p(s) + X_2(s)p(s) + \cdots + X_n(s)p(s)]$

$= \sum_{s \in S} X_1(s)p(s) + \cdots + \sum_{s \in S} X_n(s)p(s)$

$= E[X_1] + \cdots + E[X_n] = \sum_{i=1}^{n} E[X_i]$

称为期望的线性性质 (linearity of expectation)

特别地有期望线性性质并不要求随机变量之间相互独立

Probability and Statistics - P73

分布函数性质，对于随机变量X的分布函数F，即有 $F: R \to [0,1]$

令 $F(b)$ 表示随机变量X取值小于或等于实数b的概率

即有 $F(b) = Pr\{X \leq b\}$

则随机变量X的分布函数F具有性质

① 分布函数F是非降函数，即对于任意 $a<b$，有 $F(a) \leq F(b)$

② 当b趋向于 $+\infty$ 时，$F(b)=1$，即 $\lim_{b\to+\infty} F(b) = 1$

③ 当b趋向于 $-\infty$ 时，$F(b)=0$，即 $\lim_{b\to-\infty} F(b) = 0$

④ 分布函数F是右连续的，即对于任意 $b\in R$ 和单调递减收敛于b的无穷序列 $\{b_n\}$

则有 $\lim_{n\to+\infty} F(b_n) = F(b)$

证明过程有，对于性质①，有 $F(a) = Pr\{X \leq a\}$，$F(b) = Pr\{X \leq b\}$

则当 $a<b$ 时，事件 $\{X\leq a\}$ 包含于事件 $\{X \leq b\}$

于是可知 $Pr\{X\leq a\} \leq Pr\{X \leq b\}$，即 $F(a) \leq F(b)$

对于性质②，③，④由于概率的连续属性而成立

考虑单调递增的无穷序列 $\{b_n\}$，

且当 $n\to+\infty$ 时，$b_n \to +\infty$

于是事件序列 $\{X\leq b_n\}$ 为一个递增事件序列，且有 $\bigcup_{i=1}^{\infty}\{X\leq b_n\} = \{X<+\infty\}$

即 $\lim_{n\to+\infty} Pr\{X\leq b_n\} = Pr\{X<+\infty\} = 1$，即性质②成立

考虑单调递减的无穷序列 $\{b_n\}$

且当 $n\to+\infty$ 时，$b_n \to -\infty$

于是事件序列 $\{X\leq b_n\}$ 为一个递减事件序列，且有 $\bigcap_{i=1}^{\infty}\{X\leq b_n\} = \{X<-\infty\}$

即 $\lim_{n\to+\infty} Pr\{X\leq b_n\} = Pr\{X<-\infty\} = 0$，即性质③成立

考虑单调递减且收敛于b的无穷序列 $\{b_n\}$

即 $\forall_{\varepsilon>0} \exists M\in N^+ \ (n\geq M) \to (b_n - b < \varepsilon)$，即当 $n\to+\infty$ 时，$b_n \to b^+$

于是事件序列 $\{X\leq b_n\}$ 为一个递减事件序列，且有 $\bigcap_{i=1}^{\infty}\{X\leq b_n\} = \{X\leq b\}$

即 $\lim_{n\to+\infty} Pr\{X\leq b_n\} = Pr\{X\leq b\} = F(b)$，即性质④成立

注意对于随机变量X以及分布函数 $F: R\to[0,1]$

任意关于随机变量X的概率问题，均可以转换为对其分布函数值的计算

即对于任意 $a<b$，考虑事件 $\{a<X\leq b\}$

由于事件 $\{X\leq b\} = \{X\leq a\} \cup \{a<X\leq b\}$

于是有 $Pr\{X\leq b\} = Pr\{X\leq a\} + Pr\{a<X\leq b\}$

即 $Pr\{a<X\leq b\} = Pr\{X\leq b\} - Pr\{X\leq a\} = F(b) - F(a)$