確率分布 1

統計力学では直接使う場面はあまり出てきませんが、離散型の確率分布を見ます。簡単に言えば、離散型は確率変数が離散的である場合、確率分布は確率を表す関数のことです。
ここでは二項分布、超幾何分布、ポアソン分布を扱います。導出と平均、分散を求めています。
具体的な例は扱っていません。

・二項分布
・超幾何分布
・ポアソン分布

先に平均と分散の記号を定義をしておきます。確率 P_i で a_i となる量の平均は

$$< a >= \sum_i a_i P_i \quad (\sum_i P_i = 1)$$

i は a_i の個数に対応する範囲での和です。確率の話をするときの表記では、確率変数 X に対する平均は

$$E(X) = \sum_i a_i P_i(X = a_i)$$

のように書かれます。分散は

$$< (a- < a >)^2 >= \sum_i (a_i - < a >)^2 P_i \quad (< (a- < a >)^2 >=< a^2 > - < a >^2)$$

これは

$$V(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2$$

と書かれます。

- 二項分布

  結果が 2 択で出力される対象を考えます。その 2 択は成功なら S、失敗なら F として出力されるとします (例えば、コインを投げて表が出れば成功、裏なら失敗)。これを n 回繰り返します。そのとき、それぞれの結果が別の結果に影響を与えないとします (2 回目が S だからといって、次の結果が F になりやすいといったことは起きない)。なので、S が出る確率を p、F が出る確率を $q = 1 - p$ と出来ます。簡単に言えば、結果が S か F しか出ない独立な試行を n 回繰り返すということです。このとき、S が k 回出る確率を $P_n(k)$ と書きます。確率変数 X は S の回数 k です。

  例として、4 回試行し、その結果が S が 3 回、F が 1 回だったとし、この確率を $P_4(3)$ と書くことにします (添え字の 4 は試行回数、3 は S の回数。F の回数は $4 - 3 = 1$)。結果としては F, S, S, S、S, F, S, S、S, S, F, S、S, S, S, F の 4 通りがあり、それぞれの確率は全て

$$pppq = p^3 q$$

  となるので (確率は独立なので単純に積を取ればいい)、$P_4(3)$ は

$$P_4(3) = 4p^3 q$$

  となります。同様にしていけば

$$P_4(k=1) = 4pq^3$$
$$P_4(k=2) = 6p^2q^2$$
$$P_4(k=3) = 4p^3q$$
$$P_4(k=4) = p^4$$
$$P_4(k=0) = q^4$$

これらの係数は、明らかにそれぞれの S, F の組み合わせの数なので

$$_nC_k = \begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{(n-k)!k!}$$

に対応しています。実際に

$$P_4(k=2) = {}_4C_2p^2q^2 = \frac{4!}{(4-2)!2!}p^3q = 6p^2q^2$$
$$P_4(k=3) = {}_4C_3p^3q = \frac{4!}{(4-3)!3!}p^3q = 4p^3q$$

となっています。そして、さらに p^k, q^{4-k} となっていることも分かります。つまり、今の結果はまとめて

$$P_4(k) = {}_4C_kp^kq^{4-k} = {}_4C_kp^k(1-p)^{4-k}$$

と書くことが出来ます。これから、n 回の試行における確率に一般化できて

$$P_n(k) = {}_nC_kp^kq^{n-k} = {}_nC_kp^k(1-p)^{n-k}$$

となります。この確率分布を二項分布 (binomial distribution) と言い、今の場合では試行回数 n 回において S の回数が k となる確率が二項分布に従っています (確率変数 X が二項分布に従う)。$n=1$ のときはベルヌーイ分布 (Bernoulli distribution) と呼ばれます。確率分布のグラフは、横軸を k(最大値は試行回数 n)、縦軸を $P_n(k)$ として書かれます。

また、二項というのは $_nC_k$ が二項定理

$$(x+y)^n = \sum_{k=0}^{n} {}_nC_kx^ky^{n-k} = \sum_{k=0}^{n} \frac{n!}{(n-k)!k!}x^ky^{n-k}$$

での二項係数 $_nC_k$ と同じことからです。

$P_n(k)$ の k に対する和を取れば、二項定理と $p+q=1$ から

$$\sum_{k=0}^{n} P_n(k) = \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} p^k q^{n-k} = (p+q)^n = 1$$

となって、確率が規格化されていることが確かめられます (場合の数から導出しているので規格化されていないとおかしい)。

S の個数 k の平均 $<k>$ を求めます。k となる確率は $P_n(k)$ なので、平均の定義から

$$<k>= \sum_{k=0}^{n} kP_n(k) = \sum_{k=0}^{n} k\frac{n!}{(n-k)!k!} p^k q^{n-k}$$

これをどう変形すればいいかを分かりやすくするために、$n=4$ のときを見ておきます。$n=4$ とすれば、単純な変形と、$p+q=1$ であることから

$$\begin{aligned}
\sum_{k=1}^{4} kP_4(k) &= 1\times P_4(1) + 2\times P_4(2) + 3\times P_4(3) + 4\times P_4(4) \\
&= 1\times 4pq^3 + 2\times 6p^2q^2 + 3\times 4p^3q + 4\times p^4 \\
&= 4pq^3 + 12p^2q^2 + 12p^3q + 4p^4 \\
&= 4p(q^3 + 3pq^2 + 3p^2q + p^3) \\
&= 4p(p+q)^3 \\
&= 4p
\end{aligned}$$

と出来ます。一行目の右辺は、4 回試行して S が 1 回のときが第一項、S が 2 回のときが第二項、$\cdots$となっています。これから、4 回試行したとき、S となる回数の平均を求めているのがはっきりすると思います。その平均が $4p$ となっているのも分かりやすい結果です。例えば、表 (S) と裏 (F) が確率 1/2 で出るコインを 4 回投げれば、表 (S) が 2 回は出るという感覚的な結果と同じです (表が出る平均が 2 回)。

この $n=4$ での場合から、n を一般化したときは二項定理を使えばいいことが予想できます。なので、$n=4$ での変形の仕方を真似ることで、二項定理の形になるように変形します。まず、$n=4$ では $4p$ を外に出すことで $(p+q)^3$ と出来たので

$$<k>= \sum_{k=1}^{n} k\frac{n!}{(n-k)!k!} p^k q^{n-k} = np\sum_{k=1}^{n} k\frac{(n-1)!}{(n-k)!k!} p^{k-1} q^{n-k}$$

$k=0$ では 0 なので、$k=1$ から和を取るようにしています。p^{k-1} になっていることと、浮いている k が邪魔なので

$$np\sum_{k=1}^{n} k\frac{(n-1)!}{(n-k)!k!} p^{k-1} q^{n-k} = np\sum_{k=1}^{n} k\frac{1}{k}\frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} q^{n-k} = np\sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} q^{n-k}$$

後は二項定理の形になるように階乗部分を書き換えて

$$
\begin{aligned}
<k>&= np\sum_{k=1}^{n}\frac{(n-1)!}{(n-k)!(k-1)!}p^{k-1}q^{n-k}\\
&= np\sum_{k=1}^{n}\frac{(n-1)!}{(n-1-(k-1))!(k-1)!}p^{k-1}q^{n-k}\\
&= np\sum_{k=1}^{n}\frac{(n-1)!}{(n-1-(k-1))!(k-1)!}p^{k-1}q^{n-1-(k-1)}\\
&= np\sum_{l=0}^{n-1}\frac{(n-1)!}{(n-1-l)!l!}p^{l}q^{n-1-l}\qquad (l=k-1)\\
&= np\sum_{l=0}^{n-1}{}_{n-1}C_l p^l q^{n-1-l}
\end{aligned}
$$

これと二項定理を比較すれば

$$
\sum_{l=0}^{n-1}{}_{n-1}C_l p^l q^{n-1-l}=(p+q)^{n-1}
$$

と出来ます。そして、$p+q=1$ なので、平均は

$$
<k>= np\sum_{l=0}^{n-1}{}_{n-1}C_l p^l q^{n-1-l}=np(p+q)^{n-1}=np
$$

となります。というわけで、二項分布での平均は試行回数と S となる確率の積という簡単な形になります。

次に分散を求めます。分散は

$$
<(k-<k>)^2>=<k^2>-<k>^2
$$

なので、分散を求めるには k^2 の平均が必要になります。これは平均を求めたのと同様に求められます。平均の式に入れて

$$
\begin{aligned}
<k^2>=\sum_{k=1}^{n}k^2\frac{n!}{(n-k)!k!}p^kq^{n-k} &= \sum_{k=1}^{n}k^2\frac{1}{k}\frac{n!}{(n-k)!(k-1)!}p^kq^{n-k}\\
&= \sum_{k=1}^{n}k\frac{n!}{(n-k)!(k-1)!}p^kq^{n-k}\\
&= \sum_{l=0}^{n-1}(l+1)\frac{n!}{(n-(l+1))!l!}p^{l+1}q^{n-(l+1)}\qquad (l=k-1)\\
&= \sum_{l=0}^{n-1}l\frac{n!}{(n-(l+1))!l!}p^{l+1}q^{n-(l+1)}+\sum_{l=0}^{n-1}\frac{n!}{(n-(l+1))!l!}p^{l+1}q^{n-(l+1)}
\end{aligned}
$$

第一項は

$$
\begin{aligned}
\sum_{l=0}^{n-1} l\frac{n!}{(n-(l+1))!l!}p^{l+1}q^{n-(l+1)} &= \sum_{l=1}^{n-1} l\frac{n!}{(n-(l+1))!l!}p^{l+1}q^{n-(l+1)} \\
&= \sum_{l=1}^{n-1} \frac{n!}{(n-1-l)!(l-1)!}p^{l+1}q^{n-1-l} \\
&= \sum_{l=1}^{n-1} \frac{n!}{(n-1-(l-1)-1)!(l-1)!}p^{l+1}q^{n-1-l} \\
&= \sum_{l=1}^{n-1} \frac{n!}{(n-2-(l-1))!(l-1)!}p^{l+1}q^{n-1-l} \\
&= p^2\sum_{l=1}^{n-1} \frac{n!}{(n-2-(l-1))!(l-1)!}p^{l-1}q^{n-1-l} \\
&= p^2\sum_{j=0}^{n-2} \frac{n!}{(n-2-j)!j!}p^{j}q^{n-1-(j+1)} \qquad (j=l-1) \\
&= n(n-1)p^2\sum_{j=0}^{n-2} \frac{(n-2)!}{(n-2-j)!j!}p^{j}q^{n-2-j} \\
&= n(n-1)p^2(p+q)^{n-2} \\
&= n(n-1)p^2
\end{aligned}
$$

第二項は

$$
\begin{aligned}
\sum_{l=0}^{n-1} \frac{n!}{(n-(l+1))!l!}p^{l+1}q^{n-(l+1)} &= \sum_{l=0}^{n-1} \frac{n!}{(n-1-l)!l!}p^{l+1}q^{n-1-l} \\
&= np\sum_{l=0}^{n-1} \frac{(n-1)!}{(n-1-l)!l!}p^{l}q^{n-1-l} \\
&= np(p+q)^{n-1} \\
&= np
\end{aligned}
$$

よって

$$
<k^2>= n(n-1)p^2 + np
$$

というわけで、分散は

$$
\begin{aligned}
<(k-<k>)^2>= & <k^2>-<k>^2 \\
&= n(n-1)p^2+np-n^2p^2 \\
&= -np^2+np \\
&= np(1-p)
\end{aligned}
$$

となります。

- 超幾何分布

　N 個の区別可能な対象を用意します。これのうち M 個が S という性質を持ち、$N-M$ 個が F という性質を持っているとします (性質 S,F はなんでもよく、二項分布のときのように成功、失敗でもいい)。このとき、N 個の中から n 個の対象を取り出したとき、性質 S を持ったものがいくつ含まれているかの確率を考えます。ただし、N 個の中から取り出すとき、取り出したものは元に戻さないとします (非復元抽出)。二項分布は取り出したら元に戻す (復元抽出) とした場合です (例えば、ボールを戻せばボールを取り出す確率は変化しないので、コインの表裏が出る確率と同じようになる)。

　というわけで、ボールが N 個あり、そのうち白色が M 個、黒色が $N-M$ 個あるとして、ボールを n 個取り出したとき (ボールは元に戻さない)、白色のボールがいくつ含まれているかといった話になります。この例を使って具体的な場合をまずは考えていきます。$N=20$、$M=14$、$n=5$ とします。このとき、取り出した 5 個のボールの内 2 個が白色だとします。このときの場合の数は、白色を $M=14$ 個の中から 2 個、黒色を $N-M=6$ 個の中から 3 個選んだときの組み合わせなので

$$
{}_{14}C_2\times{}_6C_3=\frac{14!}{(14-2)!2!}\frac{6!}{(6-3)!3!}
$$

として求まります。そして、20 個のボールから 5 個取り出すときの場合の数は

$$
{}_{20}C_5=\frac{20!}{(20-5)!5!}
$$

なので、取り出した 5 個のボールの内 2 個が白色となる確率は

$$
P_2(N=20,M=14,n=5)=\frac{{}_{14}C_2\times{}_6C_3}{{}_{20}C_5}
$$

となります。

　これを一般化すれば

$$
P_k(N,M,n)=\frac{{}_MC_k\times{}_{N-M}C_{n-k}}{{}_NC_n}=\frac{\begin{pmatrix} M \\ k \end{pmatrix}\begin{pmatrix} N-M \\ n-k \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}
$$

このとき、取り出された白色のボールの数 k に対しては制限があります。

　まず、k の最小値は全て黒色のボールだった場合での $k=0$ ですが、0 ではいけない場合があります。これは $n>N-M$ のときに起こります。なぜなら、$N-M$ は黒色のボールの総数なので、その総数よりも取り出したボールの数 n が大きいとき、黒色のボール全てを取り出したとしても $k=0$ では n 個にならないからです。なので、$n>N-M$ では白色のボールの個数 k の最小値は $k=0$ とはならずに、$k=n-(N-M)$

個が最小になります。よって、$n > N - M$ のとき $k = n - (N - M)$、そうでなければ $k = 0$ となります。このことは

$$\max(0, n - N + M) \leq k$$

と書かれます。$\max(a, b)$ は a, b の大きい方という意味です。このことは ${}_{N-M}C_{n-k}$ の定義からも分かります。これは、$N - M < n - k$ では 0 と定義されているので

$$k \geq n - N + M$$

でなければいけないことになり、同じ制限となっています。

今度は最大値を見てみます。k が最大になるのは取り出したボールが全て白色のときなので、$k = n$ です。しかし、当然白色のボールの総数 M を超えることは出来ないので、$n > M$ での k の最大値は $k = M$ です。よって、k の最大値は、$n > M$ では $k = M$、そうでなければ $k = n$ となります。このことは

$$k \leq \min(n, M)$$

と書かれ、$\min(a, b)$ は a, b の小さい方という意味です。この場合も、${}_MC_k$ において $M < k$ では 0 になってしまうことからも出てきます。

これらをまとめて、k の取れる範囲は

$$\max(0, n - N + M) \leq k \leq \min(n, M)$$

と書かれます。

というわけで、N 個の中に性質 S を持つものが M 個、F を持つものが $N - M$ 個入っているときに、その中から非復元抽出で n 個取り出したとき、S が k 個含まれている確率を表す確率分布は

$$P_k(N, M, n) = \frac{{}_MC_k \times {}_{N-M}C_{n-k}}{{}_NC_n} = \frac{\begin{pmatrix} M \\ k \end{pmatrix}\begin{pmatrix} N - M \\ n - k \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

$$\max(0, n - N + M) \leq k \leq \min(n, M)$$

となり、これを超幾何分布 (hypergeometric distribution) と言います。このときの確率変数 X は S の個数 k です。

k に対する和を取ると

$$\begin{aligned}
\sum_{k=0}^{n} P_k(N, M, n) = \sum_{k=0}^{n} \frac{{}_MC_k \times {}_{N-M}C_{n-k}}{{}_NC_n} &= \frac{1}{{}_NC_n} \sum_{k=0}^{n} \begin{pmatrix} M \\ k \end{pmatrix}\begin{pmatrix} N - M \\ n - k \end{pmatrix} \\
&= \frac{1}{{}_NC_n} \begin{pmatrix} N \\ n \end{pmatrix} \\
&= \frac{{}_NC_n}{{}_NC_n} \\
&= 1
\end{aligned}$$

となるので、確率は規格化されています。証明は省きますが、途中で

$$\sum_{k=0}^{n}\begin{pmatrix} m \\ k \end{pmatrix}\begin{pmatrix} N-m \\ n-k \end{pmatrix}=\begin{pmatrix} N \\ n \end{pmatrix} \tag{1}$$

を使っています。

超幾何分布での平均と分散を求めます。平均は、(1) が使えるように変形していくと

$$\begin{aligned}
< k >=\sum_{k=0}^{n}kP_k(N,M,n) &= \sum_{k=0}^{n}k\frac{{}_MC_k\times{}_{N-M}C_{n-k}}{{}_NC_n} \\
&= \frac{1}{{}_NC_n}\sum_{k=1}^{n}k\frac{M!}{(M-k)!k!}{}_{N-M}C_{n-k} \\
&= M\frac{1}{{}_NC_n}\sum_{k=1}^{n}\frac{(M-1)!}{(M-1-(k-1))!(k-1)!}\begin{pmatrix} N-M \\ n-k \end{pmatrix} \\
&= M\frac{1}{{}_NC_n}\sum_{k=1}^{n}\frac{(M-1)!}{(M-1-(k-1))!(k-1)!}\begin{pmatrix} N-1-(M-1) \\ n-1-(k-1) \end{pmatrix} \\
&= M\frac{1}{{}_NC_n}\sum_{l=0}^{n-1}\frac{(M-1)!}{(M-1-l)!l!}\begin{pmatrix} N-1-(M-1) \\ n-1-l \end{pmatrix} \qquad (l=k-1) \\
&= M\frac{1}{{}_NC_n}\sum_{l=0}^{n-1}\begin{pmatrix} M-1 \\ l \end{pmatrix}\begin{pmatrix} N-1-(M-1) \\ n-1-l \end{pmatrix}
\end{aligned}$$

(1) を使えば

$$\sum_{l=0}^{n-1}\begin{pmatrix} M-1 \\ l \end{pmatrix}\begin{pmatrix} N-1-(M-1) \\ n-1-l \end{pmatrix}=\begin{pmatrix} N-1 \\ n-1 \end{pmatrix}$$

となるので、平均は

$$\begin{aligned}
< k >=M\frac{1}{{}_NC_n}\begin{pmatrix} N-1 \\ n-1 \end{pmatrix} &= M\frac{(N-n)!n!}{N!}\frac{(N-1)!}{(N-1-(n-1))!(n-1)!} \\
&= M\frac{(N-n)!n!}{N!}\frac{(N-1)!}{(N-n)!(n-1)!} \\
&= M\frac{(N-1)!}{N!}\frac{n!}{(n-1)!} \\
&= \frac{nM}{N}
\end{aligned}$$

と求まります。階乗部分は $n!/(n-1)!=n$ のようになっていて、丁寧に書けば

$$\frac{(N-1)!}{N!}\frac{n!}{(n-1)!}=\frac{(N-1)(N-2)\cdots 1}{N(N-1)\cdots 1}\frac{n(n-1)\cdots 1}{(n-1)(n-2)\cdots 1}=\frac{1}{N}n$$

ということです。

分散は二項分布のときもそうなんですが、$<k^2>$ はまともに計算するより、$<k(k-1)>$ から求めた方が簡単です。$<k^2>$ は

$$<k^2>=<k(k-1)>+<k> \qquad (<k(k-1)>=<(k^2-k)>=<k^2>-<k>)$$

で求められるために、階乗部分の変形が楽になります。$<k(k-1)>$ は

$$<k(k-1)>=\sum_{k=0}^{n}k(k-1)P_k(N,M,n)=\sum_{k=0}^{n}k(k-1)\frac{{}_MC_k\times {}_{N-M}C_{n-k}}{{}_NC_n}$$

この場合は $k\geq 2$ のときに 0 でなくなるので

$$\begin{aligned}
<k(k-1)>= &\sum_{k=2}^{n}k(k-1)\frac{{}_MC_k\times {}_{N-M}C_{n-k}}{{}_NC_n}\\
&=\frac{1}{{}_NC_n}\sum_{k=2}^{n}k(k-1)\frac{M!}{(M-k)!k!}{}_{N-M}C_{n-k}\\
&=\frac{1}{{}_NC_n}\sum_{k=2}^{n}k(k-1)\frac{1}{k}\frac{1}{k-1}\frac{M!}{(M-k)!(k-2)!}{}_{N-M}C_{n-k}\\
&=\frac{M(M-1)}{{}_NC_n}\sum_{k=2}^{n}\frac{(M-2)!}{(M-k)!(k-2)!}{}_{N-M}C_{n-k}\\
&=\frac{M(M-1)}{{}_NC_n}\sum_{k=2}^{n}\frac{(M-2)!}{(M-k)!(k-2)!}\begin{pmatrix}N-2-(M-2)\\ n-2-(k-2)\end{pmatrix}\\
&=\frac{M(M-1)}{{}_NC_n}\sum_{l=0}^{n-2}\frac{(M-2)!}{(M-2-l)!l!}\begin{pmatrix}N-2-(M-2)\\ n-2-l\end{pmatrix}\qquad (l=k-2)\\
&=\frac{M(M-1)}{{}_NC_n}\sum_{l=0}^{n-2}\begin{pmatrix}M-2\\ l\end{pmatrix}\begin{pmatrix}N-2-(M-2)\\ n-2-l\end{pmatrix}
\end{aligned}$$

このときも (1) から

$$\sum_{l=0}^{n-2}\begin{pmatrix}M-2\\ l\end{pmatrix}\begin{pmatrix}N-2-(M-2)\\ n-2-l\end{pmatrix}=\begin{pmatrix}N-2\\ n-2\end{pmatrix}$$

となるので

$$<k(k-1)>=M(M-1)\frac{(N-n)!n!}{N!}\frac{(N-2)!}{(N-n)!(n-2)!}=M(M-1)\frac{n!}{(n-2)!}\frac{(N-2)!}{N!}=M(M-1)\frac{n(n-1)}{N(N-1)}$$

階乗部分は $n!/(n-2)! = n(n-1)$ です。
　$<k(k-1)>$ から、$<k^2>$ は

$$<k^2>=<k(k-1)>+<k>= nM(n-1)\frac{M-1}{N(N-1)}+\frac{nM}{N}=\frac{nM}{N}((n-1)\frac{M-1}{N-1}+1)$$

よって、分散は

$$\begin{aligned}<k^2>-<k>^2 &= \frac{nM}{N}((n-1)\frac{M-1}{N-1}+1)-\frac{n^2M^2}{N^2}\\
&= \frac{nM}{N}((n-1)\frac{M-1}{N-1}+1-\frac{nM}{N})\\
&= \frac{nM}{N}\frac{N(n-1)(M-1)+N(N-1)-nM(N-1)}{N(N-1)}\\
&= \frac{nM}{N}\frac{nMN-nN-MN+N+N^2-N-nMN+nM}{N(N-1)}\\
&= \frac{nM}{N}\frac{N^2-nN-MN+nM}{N(N-1)}\\
&= \frac{nM}{N}\frac{(N-n)(N-M)}{N(N-1)}\end{aligned}$$

と求まります。

　超幾何分布の極限として二項分布を出せます。二項分布は超幾何関数を復元抽出にしたものなので、M, N を無限大に持っていきます。そうすれば、M の中から 1 個取り出してもボールを取り出す確率は一定とすることができます。つまり、$M, N\to\infty$ とし、白のボールを取り出す確率 $p = M/N$ は一定とします。
　まず、超幾何分布を

$$\begin{aligned}P_k(N,M,n) &= \frac{\begin{pmatrix}M\\k\end{pmatrix}\begin{pmatrix}N-M\\n-k\end{pmatrix}}{\begin{pmatrix}N\\n\end{pmatrix}}\\
&= \frac{M!}{(M-k)!k!}\frac{(N-M)!}{(N-M-(n-k))!(n-k)!}(\frac{N!}{(N-n)!n!})^{-1}\\
&= \frac{M!}{(M-k)!k!}\frac{(N-M)!}{(N-M-(n-k))!(n-k)!}\frac{(N-n)!n!}{N!}\\
&= \frac{n!}{(n-k)!k!}\frac{M!}{(M-k)!}\frac{(N-M)!}{(N-M-n+k)!}\frac{(N-n)!}{N!}\\
&= {}_nC_k(M(M-1)\cdots(M-k+1))\\
&\quad \times((N-M)(N-M+1)\cdots(N-M-n+k+1))\frac{1}{N(N-1)\cdots(N-n+1)}\end{aligned}$$

と変形します。取り出す白のボールの数 k は取り出すボールの総数 n より小さいので

$$\frac{1}{N(N-1)\cdots(N-n+1)}=\frac{1}{N(N-1)\cdots(N-k+1)}\frac{1}{(N-k)(N-k-1)\cdots(N-n+1)}$$

とすれば、

$$\frac{M(M-1)\cdots(M-k+1)}{N(N-1)\cdots(N-k+1)}\ ,\ \frac{(N-M)(N-M+1)\cdots(N-M-n+k+1)}{(N-k)(N-k-1)\cdots(N-n+1)}$$

というのが出てきます。これらの M,N の無限大の極限は、一番大きなオーダを持つ M,N の寄与だけを取り出せばいいので

$$\lim_{M,N\to\infty}\frac{M(M-1)\cdots(M-k+1)}{N(N-1)\cdots(N-k+1)}\simeq\lim_{M,N\to\infty}\frac{M^k}{N^k}=p^k$$

と

$$\begin{aligned}\lim_{M,N\to\infty}\frac{(N-M)(N-M+1)\cdots(N-M-n+k+1)}{(N-k)(N-k-1)\cdots(N-n+1)}&\simeq\lim_{M,N\to\infty}\frac{(N-M)^{n-k}}{N^{n-k}}\\&=\lim_{M,N\to\infty}(1-\frac{M}{N})^{n-k}\\&=(1-p)^{n-k}\end{aligned}$$

となります。それぞれ、$M-(M-k+1)+1=k$ 個の積、$N-M-(N-M-n+k+1)+1=n-k$ 個の積なので、k 乗と $n-k$ 乗だけを取り出しています。よって

$$\lim_{M,N\to\infty}P_k(N,M,n)={}_nC_kp^k(1-p)^{n-k}$$

白のボールを取り出すことが成功 S に対応し、その確率は p なので、二項分布となります。

- ポアソン分布

　ポアソン分布の典型的な導出としては、ポアソン過程と二項分布の極限から求める方法があります。先にポアソン過程から求めて、その後に二項分布の極限として求めます。

　ポアソン過程の前に逆算したような方法でポアソン分布を出しておきます。ある時間間隔 t の間に起きる事象を考え、同じ間隔 t で事象が起きた回数を繰り返し数えます。間隔は時間でなく距離とかの適当な間隔でもいいです。事象が起きる確率は、時間間隔において一定で、事象が何回起きていようとも同じだとします。これは例えば、10 分の間にすれ違う人の数を 5 回計測したとき、0 人だったのが 1 回、1 人だったのが 1 回、2 人だったのが 2 回、3 人だったのが 1 回というようなことです。

　結果が成功 S と失敗 F で出力される対象を用意します (二項分布と対応させやすいように S,F にします)。時間間隔 t において S が出力される回数を繰り返し数えます。その結果は、時間間隔 t において、S は 0 回だったのが x_0 回、1 回だったのが x_1 回、…、m 回だったのが x_m 回だとします。そして、このときの S の回数の平均は $\lambda>0$ だとし ($\lambda=(0\times x_0+1\times x_1+\cdots+m\times x_m)/(x_0+x_1+\cdots+x_m)$)、分散も λ だとします。t の間に S の回数が k 回となる確率 $P(k)$ は

$$P(k)=\frac{x_k}{x_0+x_1+\cdots+x_m}\tag{2}$$

となっています。

今は k の平均 $<k>$ と分散 $<k^2>-<k>^2$ は λ になるとしています。これを満たすために、S が k 回となった回数 x_k に対して

$$x_1=\lambda x_0\ ,\ x_2=\frac{\lambda}{2}x_1=\frac{\lambda^2}{2!}x_0\ ,\ x_3=\frac{\lambda}{3}x_1=\frac{\lambda^3}{3!}x_0\ ,\ \dots\ ,\ x_m=\frac{\lambda^m}{m!}x_0 \tag{3}$$

という関係を持つことを要求します。これによって平均は

$$\begin{aligned}
<k>&=\frac{1}{x_0(1+\lambda+\lambda^2/2!+\cdots+\lambda^m/m!)}(1\times\lambda x_0+2\frac{\lambda^2}{2!}x_0+3\frac{\lambda^3}{3!}x_0+\cdots+m\frac{\lambda^m}{m!}x_0)\\
&=\frac{1}{1+\lambda+\lambda^2/2!+\cdots+\lambda^m/m!}(\lambda+\lambda^2+\frac{1}{2!}\lambda^3+\cdots+\frac{\lambda^m}{(m-1)!})\\
&=\lambda\frac{1}{(1+\lambda+\lambda^2/2!+\cdots+\lambda^m/m!)}(1+\lambda+\frac{1}{2!}\lambda^2+\cdots+\frac{\lambda^{m-1}}{(m-1)!})
\end{aligned}$$

このとき、$m-1$ を $m-1\simeq m$ と近似できる程度に m が大きければ平均を

$$<k>\simeq\lambda$$

と出来ます。また、極限

$$\lim_{n\to\infty}\frac{x^n}{n!}=0\qquad(x>0)$$

から、n が十分大きければ、この形の項はほぼ 0 と見なせます。

$<k^2>$ は m が十分大きければ

$$\begin{aligned}
<k^2>&=\frac{1}{x_0(1+\lambda+\lambda^2/2!+\cdots+\lambda^m/m!)}(1^2\times\lambda x_0+2^2\frac{\lambda^2}{2!}x_0+3^2\frac{\lambda^3}{3!}x_0+\cdots+m^2\frac{\lambda^m}{m!}x_0)\\
&=\frac{1}{1+\lambda+\lambda^2/2!+\cdots+\lambda^m/m!}(\lambda+2\lambda^2+\frac{3}{2!}\lambda^3+\cdots+\frac{m\lambda^m}{(m-1)!})\\
&=\frac{1}{1+\lambda+\lambda^2/2!+\cdots+\lambda^m/m!}\lambda(1+2\lambda+\frac{3}{2!}\lambda^2+\cdots+\frac{m\lambda^{m-1}}{(m-1)!})\\
&\simeq\lambda(1+\lambda)\frac{1}{(1+\lambda+\lambda^2/2!+\cdots+\lambda^m/m!)}(1+\lambda+\frac{1}{2!}\lambda^2+\cdots+\frac{\lambda^m}{m!})\\
&\simeq\lambda(1+\lambda)
\end{aligned}$$

$\simeq$ は m を十分大きく取ることでの近似で、下から二行目では

$$
\begin{aligned}
(1+\lambda)\sum_{k=0}^{n}\frac{1}{k!}\lambda^k =& \sum_{k=0}^{n}\frac{1}{k!}\lambda^k+\sum_{k=0}^{n}\frac{1}{k!}\lambda^{k+1} \\
=& \sum_{k=0}^{n}\frac{1}{k!}\lambda^k+\sum_{l=1}^{n+1}\frac{1}{(l-1)!}\lambda^{l} \quad (l=k+1) \\
=& 1+\sum_{k=1}^{n}\frac{1}{k!}\lambda^k+\sum_{l=1}^{n}\frac{1}{(l-1)!}\lambda^{l}+\frac{1}{n!}\lambda^{n+1} \\
=& 1+\sum_{k=1}^{n}(\frac{1}{k!}+\frac{1}{(k-1)!})\lambda^k+\frac{1}{n!}\lambda^{n+1} \\
=& 1+\sum_{k=1}^{n}\frac{1+k}{k!}\lambda^k+\frac{1}{n!}\lambda^{n+1} \quad (\frac{(k-1)!+k!}{k!(k-1)!}=\frac{(k-1)!(1+k)}{k!(k-1)!}=\frac{1+k}{k!})
\end{aligned}
$$

を使っています $(n=m-1)$。よって、分散も

$$
< k^2 > - < k >^2 \simeq \lambda(1+\lambda)-\lambda^2=\lambda
$$

と出来ます。

確率 $P(k)$ は (2) で与えているので、x_k の関係から $P(k)$ は

$$
P(1)=\frac{\lambda}{1!}P(0)\ ,\ P(2)=\frac{\lambda}{2}P(1)=\frac{\lambda^2}{2!}P(0)\ ,\ P(3)=\frac{\lambda}{3}P(2)=\frac{\lambda^3}{3!}P(0)\ ,\ \cdots\ ,\ P_m=\frac{\lambda^m}{m!}P(0)
$$

という関係を持つことになります。つまり、確率の関係も

$$
P(k)=\frac{\lambda^k}{k!}P(0) \quad (P(k)=\frac{\lambda}{k}P(k-1)) \tag{4}
$$

となっています。この関係 (4) から、確率の規格化

$$
\sum_{k=0}^{m}P_k=1
$$

は

$$
P(0)(1+\frac{\lambda}{1!}+\cdots+\frac{\lambda^m}{m!})=1
$$

となります。平均と分散を λ にするためには m を大きく取る必要があることと、これの括弧部分は exp の展開

$$
e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!}
$$

と同じ形をしていることから、m を無限大まで持っていくことにして

$$P(0)e^{\lambda} = 1$$
$$P(0) = e^{-\lambda}$$

よって、m を無限大にして、$P(0) = e^{-\lambda}$ とすることで確率の規格化が成立します。そうすると、$P(k)$ は

$$P(k) = \frac{\lambda^k}{k!}P(0) = \frac{\lambda^k}{k!}e^{-\lambda} \quad (\sum_{k=0}^{\infty} P(k) = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1)$$

となります。

この $P(k)$ によって平均と分散が λ になることを確かめておきます。平均は

$$\begin{aligned}
<k>= \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!}e^{-\lambda} &= \sum_{l=0}^{\infty}\frac{\lambda^{l+1}}{l!}e^{-\lambda} \quad (l = k-1) \\
&= \lambda\sum_{l=0}^{\infty}\frac{\lambda^l}{l!}e^{-\lambda} \\
&= \lambda e^{\lambda}e^{-\lambda} \\
&= \lambda
\end{aligned}$$

分散は、$<k(k-1)>$ は

$$\begin{aligned}
<k(k-1)>= \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} &= \sum_{k=2}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} \\
&= \sum_{k=2}^{\infty}\frac{\lambda^k}{(k-2)!}e^{-\lambda} \\
&= \sum_{l=0}^{\infty}\frac{\lambda^{l+2}}{l!}e^{-\lambda} \quad (l = k-2) \\
&= \lambda^2\sum_{l=0}^{\infty}\frac{\lambda^l}{l!}e^{-\lambda} \\
&= \lambda^2
\end{aligned}$$

となることから、

$$<k^2> - <k>^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \quad (<k^2>=<k(k-1)> + <k>= \lambda^2 + \lambda)$$

と求まります。

よって、平均と分散が λ であるとき、時間間隔 t において S が出力された回数が k 回となる確率は

$$P(k,\lambda)=\frac{\lambda^k}{k!}e^{-\lambda}\qquad(k=0,1,2,\dots)$$

と与えられます。m を無限大に取っているので、k の上限はありません。このように、間隔 t において事象が起きる回数の平均と分散が λ で与えられているとき、間隔 t で k 回事象が起きる確率を表す確率分布をポアソン分布 (Poisson distribution) と言います。調べたい事象の結果が (3) のような関係を持てばポアソン分布に近づきます。

また、時間間隔 t における S の回数の平均 λ は、単位時間あたりの平均 α を使って $\lambda=\alpha t$ とできます。これによってポアソン分布は

$$P(k,\lambda)=\frac{(at)^k}{k!}e^{-at}$$

とも書かれます。

ポアソン分布では k に上限がないので、k をどこまでも大きく取れます。しかし、極限として

$$\lim_{n\to\infty}\frac{x^n}{n!}=0\qquad(x>0)$$

があるので、k が大きい領域での確率は 0 に近づきます (ポアソン分布の形は山なり)。

今の話を数学側に寄せて見直します。まず、ポアソン過程を導入します。時間間隔 t で、ある事象が S となる回数 $k\geq 0$ を考え、その確率は $P(t,k)$ とします。このとき

(i) $t=0$ のとき確率は $P(t=0,k=0)=1,\ P(t=0,k\neq 0)=0$。

(ii) 時間間隔 $t, t+s$ における回数 k_t, k_{t+s} の差 $k_{t+s}-k_t$ は t に依存しない。

(iii) $t_1<\cdots<t_n$ での S の回数 k_i に対して $k_{i+1}-k_i$ $(i=1,2,\dots,n-1)$ は独立。

(iv) 微小間隔 Δt の間では S の回数は 1 回、もしくは 0 回。そして、1 回となる確率は Δt に比例する。

となっているならポアソン過程となります。

(i) は初期条件で、時間間隔が 0 なら S は 1 回も出力されないということを言っています (このため $k=0$ の確率は 1 で、$k\neq 0$ となる確率は 0)。

(ii) は定常増分性 (stationary increments)、(iii) は独立増分性 (independent increments) と呼ばれます。(ii) は S の回数 k は時間間隔 t にのみ依存するということ、(iii) は 2 つの時間間隔 t と t' が重なっていないとき、時間間隔 t における S の回数は時間間隔 t' での回数に影響を与えないということです。

(iv) は微小間隔において、S は最大で 1 回しか出力されないということです。1 回出力される確率は Δt に比例させるので、比例定数 α を定義します。これによって、$t+\Delta t$ と t の間に S が 1 回出力される確率は (t のとき k 回あり、$t+\Delta t$ のときに $k+1$ 回ある確率)

$$P(t+\Delta t,k+1;t,k)=\alpha\Delta t+O(\Delta t)\tag{5}$$

と書けます。確率は Δt に比例するとしますが、そこからのずれも加えることにして、それを $O(\Delta t)$ としています。$O(\Delta t)$ は

$$\lim_{\Delta t\to 0}\frac{O(\Delta t)}{\Delta t}=0$$

とします (例えば、$O(\Delta t)$ は Δt の 2 次以上のみを含んでいる)。後で分かりますが、$O(\Delta t)$ はここでの話に影響を与えないので無視してもいいです。Δt の間では S が 0 個となる確率もあるので、それは (t のとき k 回あり、$t+\Delta t$ のときに k 回ある確率)

$$P(t+\Delta t, k; t, k) = 1 - \alpha\Delta t - O(\Delta t) \tag{6}$$

と書けます (Δt の間での S の回数は 1 回か 0 回なので、1 回のときの確率を 1 から引けばいい)。

ここで、時間間隔 $t+\Delta t$ で S が k 回ある確率 $P(t+\Delta t, k)$ を考えます。この場合、(iv) から時間間隔 t での回数は $k-1$ か k で、この 2 つの場合を経由したものが $t+\Delta t$ で k となります。つまり、$t+\Delta t$ で k となるには、t で $k-1$ 回 (k 回) あり、Δt の間に 1 回 (0 回) ある必要があるので、確率 $P(t+\Delta t, k)$ はそれぞれの経由した状況の確率をかけて足したものになります。

よって、t で $k-1$ の確率を $P(t, k-1)$、t で k の確率を $P(t, k)$ とすれば、Δt の間に 1 回もしくは 0 回ある確率は (5),(6) によるので

$$\begin{aligned}
P(t+\Delta t, k) &= P(t+\Delta t, k; t, k-1)P(t, k-1) + P(t+\Delta t, k; t, k)P(t, k) \\
&= (\alpha\Delta t + O(\Delta t))P(t, k-1) + (1 - \alpha\Delta t - O(\Delta t))P(t, k) \\
&= \alpha\Delta t P(t, k-1) + (1 - \alpha\Delta t)P(t, k) + O(\Delta t)
\end{aligned}$$

これの両辺を Δt で割れば

$$\begin{aligned}
\frac{P(t+\Delta t, k)}{\Delta t} &= \alpha P(t, k-1) + \frac{P(t, k)}{\Delta t} - \alpha P(t, k) + \frac{1}{\Delta t}O(\Delta t) \\
\frac{P(t+\Delta t, k) - P(t, k)}{\Delta t} &= \alpha(P(t, k-1) - P(t, k)) + \frac{1}{\Delta t}O(\Delta t)
\end{aligned}$$

$\Delta t \to 0$ の極限を取れば、左辺は微分の定義になり、右辺第二項は消えるので ($O(\Delta t)$ は Δt の 2 次以上だから)

$$\begin{aligned}
\lim_{\Delta t \to 0} \frac{P(t+\Delta t, k) - P(t, k)}{\Delta t} &= \alpha(P(t, k-1) - P(t, k)) + \lim_{\Delta t \to 0} \frac{1}{\Delta t}O(\Delta t) \\
\frac{dP(t, k)}{dt} &= \alpha(P(t, k-1) - P(t, k))
\end{aligned}$$

見やすくするために

$$\frac{dP_k(t)}{dt} = \alpha(P_{k-1}(t) - P_k(t)) \tag{7}$$

と書くことにします。これを解きます。

差分方程式の形をしているので、$k=0$ から順番に求めていきます（差分方程式なので、(4) の形をした関係が出てくる)。$k=0$ の場合は (7) の第一項を 0 にすればいいですが、一応確率の形から出しておきます。$t+\Delta t$ で $k=0$ となる確率 $P(t+\Delta t, 0)$ は、t では $k=0$ しか経由できないので $P(t, 0)$ と (6) から

$$P(t+\Delta t, 0) = P(t+\Delta t, 0; t, 0)P(t, 0)$$

なので

$$\begin{aligned}
P(t+\Delta t,0) &= (1-\alpha\Delta t - O(\Delta t))P(t,0) \\
\lim_{\Delta t\to 0}\frac{P(t+\Delta t,0)-P(t,0)}{\Delta t} &= -\alpha P(t,0) \\
\frac{dP_0(t)}{dt} &= -\alpha P_0(t)
\end{aligned}$$

となります。これの解はすぐに

$$P_0(t) = A_0 e^{-\alpha t}$$

となっているのが分かります。定数 A_0 を決めるために (i)$P_0(0) = 1$ を使えば

$$P_0(t) = e^{-\alpha t}$$

これを (7) に入れれば $P_1(t)$ が求まります。
　$k = 1$ とした微分方程式にこの結果を入れれば

$$\begin{aligned}
\frac{dP_1(t)}{dt} &= \alpha(P_0(t) - P_1(t)) \\
\frac{dP_1(t)}{dt} + \alpha P_1(t) &= \alpha P_0(t) \\
&= \alpha e^{-\alpha t}
\end{aligned} \tag{8}$$

これは非同次の微分方程式なので、$P_1(t) = A(t)e^{-\alpha t}$ とすれば特解は

$$\begin{aligned}
\frac{d}{dt}(A(t)e^{-\alpha t}) + \alpha A(t)e^{-\alpha t} &= \alpha e^{-\alpha t} \\
e^{-\alpha t}\frac{d}{dt}A(t) - \alpha A(t)e^{-\alpha t} + \alpha A(t)e^{-\alpha t} &= \alpha e^{-\alpha t} \\
\frac{d}{dt}A(t) &= \alpha \\
A(t) &= \alpha t
\end{aligned}$$

から

$$P_1(t) = \alpha t e^{-\alpha t}$$

と求まります。同次

$$\frac{dP_1(t)}{dt} + \alpha P_1(t) = 0$$

での一般解は $Ce^{-\alpha t}$ なので (C は定数)、(8) の一般解は

$$P_1(t) = \alpha t e^{-\alpha t} + Ce^{-\alpha t}$$

ここで条件として (i)$P_1(0) = 0$ を使えば、$C = 0$ となり

$$P_1(t) = \alpha t e^{-\alpha t}$$

これを (7) に入れて $P_2(t)$ を求めます。

$P_2(t)$ での微分方程式は

$$\begin{aligned}
\frac{dP_2(t)}{dt} &= \alpha(P_1(t) - P_2(t)) \\
&= \alpha(\alpha t e^{-\alpha t} - P_2(t)) \\
\frac{dP_2(t)}{dt} + \alpha P_2(t) &= \alpha^2 t e^{-\alpha t} \qquad (9)
\end{aligned}$$

$P_2(t) = A_2(t)e^{-\alpha t}$ として入れれば

$$\begin{aligned}
e^{-\alpha t}\frac{d}{dt}A_2(t) - \alpha A_2(t)e^{-\alpha t} + \alpha A_2(t)e^{-\alpha t} &= \alpha^2 t e^{-\alpha t} \\
\frac{d}{dt}A_2(t) &= \alpha^2 t \\
A_2 &= \frac{1}{2}\alpha^2 t^2
\end{aligned}$$

同次での解は $Ce^{-\alpha t}$ なので、一般解は

$$P_2(t) = \frac{1}{2}\alpha^2 t^2 e^{-\alpha t} + Ce^{-\alpha t}$$

ここでも、(i)$P_2(0) = 0$ から $C = 0$ となり

$$P_2(t) = \frac{1}{2}\alpha^2 t^2 e^{-\alpha t}$$

となります。

ここまで来れば法則性が分かってきます。P_3 になれば、(8),(9) の右辺を見れば明らかに α^3 が出てき、t^2 の積分からは $t^3/3$ が出てくるので、$P_3(0) = 0$ の条件によって

$$P_3(t) = \frac{1}{6}\alpha^3 t^3 e^{-\alpha t}$$

となります。なので、(i)$P_0(0) = 1,\ P_k(0) = 0\ (k \neq 0)$ によって、微分方程式 (7) の解は

$$P_k(t) = \frac{(\alpha t)^k}{k!} e^{-\alpha t}$$

と求まります。これは時間間隔 t で S が k 回出力される確率を表し、ポアソン分布の形です。つまり、条件 (i),(ii),(iii),(iv)(ポアソン過程) において、時間間隔 t の間に事象が k 回起きる確率は $\lambda = \alpha t$ としたポアソン分布に従い、その平均は αt ということになります。このため、ポアソン過程の定義として、(iv) をポアソン分布に従うとすることもできます。

ちなみに

$$\begin{aligned}
\frac{dP_k(t)}{dt} &= \alpha(P_{k-1}(t) - P_k(t)) \\
\frac{dP_k(t)}{dt} + \alpha P_k(t) &= \alpha P_{k-1}(t) \\
e^{-\alpha t}\frac{d}{dt}(e^{\alpha t}P_k(t)) &= \alpha P_{k-1}(t) \\
\frac{d}{dt}(e^{\alpha t}P_k(t)) &= \alpha e^{\alpha t} P_{k-1}(t) \\
e^{\alpha t}P_k(t) &= \alpha \int_0^t dt' e^{\alpha t'} P_{k-1}(t') \\
P_k(t) &= \alpha e^{-\alpha t} \int_0^t dt'\ e^{\alpha t'} P_{k-1}(t')
\end{aligned}$$

として、$P_0(t) = e^{-\alpha t}$ から始めていっても求められます。

今度は二項分布を持ってきます。二項分布は n 回の試行で成功 S が何回出るかに対する確率、ポアソン分布は時間間隔 t で成功 S が何回あるかの確率を求めています。なので、時間間隔 t のことを無視すれば同じ確率を求めているように思えます。そして、状況で明確に異なっているのは、二項分布では S の回数 k は試行回数 n によって $k \leq n$ に制限されているのに対して、ポアソン分布では S の回数 k は無限大まで取れるようにしています。

このことから、二項分布において試行回数 n を無限大にしてみます。そのために、二項分布を

$$\begin{aligned}
P_n(k) = {}_nC_k p^k q^{n-k} &= {}_nC_k p^k (1-p)^{n-k} \\
&= \frac{n!}{(n-k)!k!} p^k q^{n-k} \\
&= \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} p^k (1-p)^{n-k} \\
&= \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \frac{1}{n^k} (np)^k (1-p)^{n-k} \\
&= \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} \frac{1}{k!} (np)^k (1-p)^{n-k}
\end{aligned}$$

と変形します。このとき、分子の積は $n-(n-k+1)+1=k$ 個によるものなので

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}=\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n}=1\times(1-\frac{1}{n})\times(1-\frac{2}{n})\times(1-\frac{k-1}{n})$$

となっています。なので、n を無限大に持っていくと

$$\lim_{n\to\infty}\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}=1$$

となることから

$$\lim_{n\to\infty}P_n(k)=\frac{1}{k!}(np)^k\lim_{n\to\infty}(1-p)^{n-k}$$

$np=\lambda$ とすれば

$$\lim_{n\to\infty}P_n(k)=\frac{\lambda^k}{k!}\lim_{n\to\infty}(1-\frac{\lambda}{n})^{n-k}$$

右辺の極限は

$$\lim_{n\to\infty}(1-\frac{\lambda}{n})^{n-k}=\lim_{n\to\infty}\frac{(1-\lambda/n)^n}{(1-\lambda/n)^k}=\lim_{n\to\infty}(1-\frac{\lambda}{n})^n$$

これは exp の極限による表現

$$e^x=\lim_{n\to\infty}(1+\frac{x}{n})^n$$

から

$$\lim_{n\to\infty}P_n(k)=\frac{\lambda^k}{k!}e^{-\lambda}$$

となり、ポアソン分布になります。つまり、ポアソン分布は、二項分布において試行回数 n を無限大にし、成功 S となる確率 $p=\lambda/n$ が小さいとした場合と同じです。このため、ポアソン分布は二項分布の極限 (近似) として使われます。二項分布とポアソン分布を見比べれば分かるように、n が大きいときはポアソン分布と近似してしまえば、計算は簡単になります。

二項分布での平均と分散に対して、$n\to\infty,\ p\to 0,\ np=\lambda$ とすることでポアソン分布の平均と分散

$$<k>=np=\lambda$$
$$<k^2>-<k>^2=np(1-p)=\lambda(1-p)\ \Rightarrow\ \lambda$$

が求められます。