III-2-3. ポアソン分布

ポアソン分布は、二項分布の期待値npを一定にして、nを無限大にした分布です。ですからpは無限小になります。このことからわかるように、ポアソン分布は、繰り返しの数が多いにもかかわらず、まれにしか起こらない現象を統計的に扱うための確率分布です。ポアソン分布は、２項分布を単純に変形することで得られるので、正規分布より拡張のプロセスを説明するのが簡単なのですが、途中に一か所だけ、自然対数の底（eのこと、数学的にはネイピア数というのが普通です。）とは何かという説明があります。ネイピア数はいろいろなところに出てくるので、きちんと理解しておく方が良いでしょう。ということで、別項を設けて説明しましたので必要なそちらを読んでください（III-3-2. ネイピア数)。

さて、ポアソン分布の説明ですがやっていることは式の変形にすぎません。ポイントはnがどんな値をとっても$np=\mu$で一定ということです。
ポアソン分布らしく、*W(k)*ではなくて、*P(k)*と書きます。ある事象が k 回起きる確率です。

$$P(k) ={}_n C_k p^k q^{(1-k)}$$

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

$$= \frac{n!}{k!\,(n-k)!} p^k (1-p)^{(n-k)}$$

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

$\frac{1}{k!}$には p も n も含まれないので、これは考える必要ないものとして次のように変形します。

$$= \frac{1}{k!} \mathrm{A}\cdot\mathrm{B}$$

$$\mathrm{A} = \frac{n!}{(n-k)!} p^k$$

$$\mathrm{B} = (1-p)^{(n-k)}$$

と二つの部分に分けて、n→∞の値を考えます。
A について

$$\mathrm{A} = \frac{n!}{(n-k)!} p^k$$

$$= \mathrm{n}(n-1)\cdots(n-k+1)p^k$$

$$= \frac{n}{n}\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-\mathrm{k}+1}{n}\right) n^k p^k$$

$$= 1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)(np)^k$$

$np=\mu$なので

$$=1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)(\mu)^k$$

カッコの中の分数はｎが大きくなれば０に近づくから、

$$\lim_{n\to\infty}A=\mu^k$$

Ｂについて

$$\mathrm{B}=(1-p)^{(n-k)}$$

ここでも$np=\mu$を使うと、

$$=\left(1-\frac{\mu}{n}\right)^{\mu\frac{n}{\mu}-k}$$

$$=\frac{\left(1-\frac{\mu}{n}\right)^{\mu\frac{n}{\mu}}}{\left(1-\frac{\mu}{n}\right)^{-k}}$$

$n\to\infty$のとき、$\left(1-\frac{\mu}{n}\right)\to 1$だから、分母は１に近づく

分子については、ネイピア数(e)が

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

であることを知っていれば、$-\frac{\mu}{n}$を$\frac{1}{n}$として、

$$=e^{-\mu}$$

となることは、簡単にわかります。問題はネイピア数を知っているかということになります。

必要ならば、III-3-2. （ネイピア数）の説明を読めばよいので、知っていることにして

$$\lim_{n\to\infty}\mathrm{B}=\frac{\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^{\mu\frac{n}{\mu}}}{\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^{-k}}$$

$$=\frac{e^{\mu}}{1}$$

$$=e^{\mu}$$

$$\lim_{n\to\infty}P(k)=\lim_{n\to\infty}{}_nC_kp^kq^{(1-k)}$$

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

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

$$=\frac{1}{k!}\lim_{n\to\infty}A\lim_{n\to\infty}B$$

$$=\frac{\mu^k e^{-\mu}}{k!}$$

となりますが、データーから μ はわからないから、その期待値としてデーターの平均値 λ をつかって

$$\mathrm{P(X = k)} = \frac{\lambda^k e^{-\lambda}}{k!}$$

式 16

式の意味は x をあることが起こる確率としたときに、それが k 回起こる確率です。なぜ、この場合に平均値の記号に λ を使うのかはよくわかりません。知りません。μ ではないという意味ならば何でもよさそうにも思いますが、慣習として λ を使います。

ポアソン分布では、平均値に依存する形で分散が変化します。２項分布だって分散が平均値によって変わります。ポアソン分布はnpが一定で、pが小さくなればnが大きくなる。nが大きくなれば分布が平均値を中心に尖ってくるのだから、当たり前だろうと言われればそれまでです。しかし、完全なポアソン分布では、分散の値は平均の値と同じになります。この性質は、データーの分散がポアソン的であるかそうでないかを考えるときに重要ですので、その証明をします。式 11 を使います。

$$V_{(k)} = E_{(k^2)} - {E_{(k)}}^2$$

確認すると、式の意味は、二乗の平均値から平均値の二乗を引くと分散になるということです。２項分布で分散を求めるときに使いました。ポアソン分布は二項分布の極限ですから、この関係が成り立ちます。

$\mathrm{E}_{(k)}$は期待値でnpですが、データーからこの値を推測すると平均値 λ です。二乗の平均値は次のように表せます。

ポアソン分布ではデーターが k である確率は、

$$\mathrm{P(X = k)} = \frac{\lambda^k e^{-\lambda}}{k!}$$

ですから

$$E_{(k^2)} = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k e^{\lambda}}{k!}$$

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

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

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

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

二項目のΣの式は k=k-1 としたときのポアソン分布の確率の総和だから 1 です。よって

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

上と同じことを繰り返します。

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

$$= \sum_{k=0}^{\infty} \frac{\lambda^k e^{\lambda}}{(k-2)!} + \lambda$$

$$= \lambda^2 \sum_{k=0}^{\infty} \frac{\lambda^{(k-2)} e^{\lambda}}{(k-2)!} + \lambda$$

二項目のΣの式は k=k-2 としたときのポアソン分布の確率の総和だから 1 です。よって

$$E_{(k^2)} = \lambda^2 + \lambda$$

$$E_{(k)} = \lambda$$

とするのだから

$$E_{(k)}{}^2 = \lambda^2$$

$$V_{(k)} = E_{(k^2)} - E_{(k)}{}^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

以上、証明終わり。