確率と統計 (0)
「確率分布の例 (第6章)」

- 担当教員： 杉山　将 (計算工学専攻)
- 居室： W8E-505
- 電子メール: sugi@cs.titech.ac.jp

- 授業のウェブサイト：
  http://sugiyama-www.cs.titech.ac.jp/~sugi/

主な離散型の確率分布

- 一様分布
- 超幾何分布
- 二項分布
- ポアソン分布
- 負の二項分布
- 幾何分布

二項分布の極限

- 二項分布において，実験の成功率が非常に低い（ p が非常に小さい）場合，めったに実験は成功しない．
- しかしいくら p が小さくても，実験の回数が非常に多い（ n が非常に大きい）場合，ある程度の回数は実験は成功するはず．
- 例：成功率が０．２%の実験を１０００回繰り返したとき，二項分布の期待値は

$$E(X) = np = 1000 \times 0.002 = 2$$

このとき，$x = 0, 1, 2, 3$ くらいの生起確率はそれほど小さくなさそう？

二項分布の極限 (続き)

- しかし, $p = 0.002$, $n = 1000$ の二項分布が $x = 2$ となる確率を計算するのは大変！

$$f(2) = {}_{1000}C_2 (0.002)^2 (0.998)^{998}$$

- ポアソンの小数の法則(Poisson's law of small numbers): $p = \lambda / n$ に対して,

$${}_nC_x p^x (1-p)^{n-x} \longrightarrow \frac{e^{-\lambda} \lambda^x}{x!}$$

$$(n \to \infty)$$

ポアソンの小数の法則

■ **証明：** $\lim_{n\to\infty} {}_nC_x \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$

$$= \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{n!}{(n-x)!n^x} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x}$$

ここで

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

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

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

が成り立つので， $\lim_{n\to\infty} {}_nC_x \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x} \longrightarrow \frac{e^{-\lambda}\lambda^x}{x!}$

ポアソン分布

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

- これを**ポアソン分布**(Poisson distribution)と呼び，$Po(\lambda)$ で表す
- $f(x)$ が確率関数であることの証明
  - $f(x) \geq 0$ は明らか
  - 指数関数の原点周りでのテーラー展開

$$e^{\lambda} = 1 + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \cdots = \sum_x \frac{\lambda^i}{x!}$$

より，

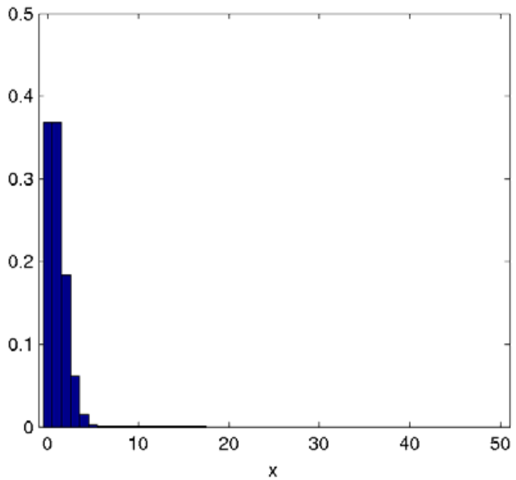
$$\sum_x f(x) = \sum_x \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_x \frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1$$

ポアソン分布の性質

- 期待値: $E(X) = \lambda$
- 分散: $V(X) = \lambda$
- 積率母関数: $M_X(t) = \exp\left\{\lambda(e^t - 1)\right\}$
- 証明は宿題
- ポアソン分布は期待値と分散が等しい

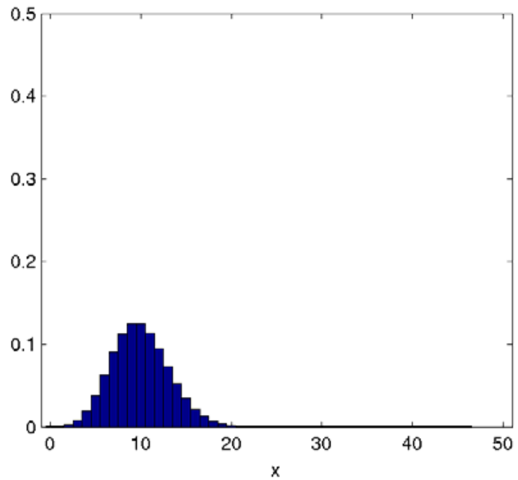
ポアソン分布の例

- 実験の平均成功回数を変えたとき

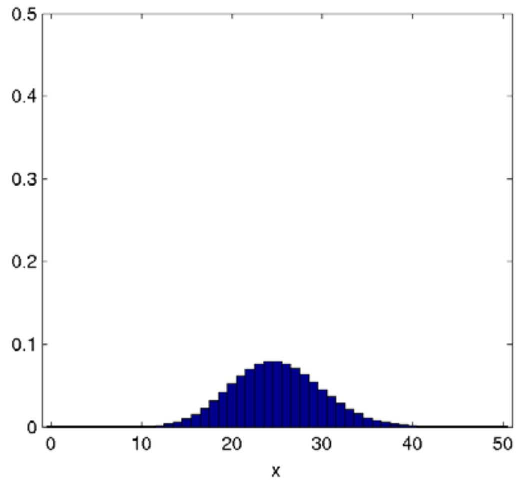


$$\lambda = 1$$

平均成功回数少ない



$$\lambda = 10$$



$$\lambda = 25$$

平均成功回数多い

負の二項分布

- ある実験の成功と失敗の確率がそれぞれ p と $1-p$
- k回目の成功を得るまでの失敗の回数を x で表せば，その確率は

$$f(x) = {}_{k+x-1}C_x p^k (1-p)^x$$

$$\text{for } x = 0, 1, 2, \ldots$$

- これを**負の二項分布**(negative binomial distribution)，または**パスカル分布**(Pascal distribution)という

負の二項係数

- 二項係数の実数への拡張

$$
{}_{a}C_x = \frac{(a)\times(a-1)\times\cdots\times(a-x+1)}{x\times(x-1)\times\cdots\times 1}
$$

- 負の二項分布の名前の由来：負の二項係数

$$
\begin{aligned}
{}_{-k}C_x &= (-1)^x \frac{k\times(k+1)\times\cdots\times(k+x-1)}{x\times(x-1)\times\cdots\times 1} \\
&= (-1)^x\,{}_{k+x-1}C_x
\end{aligned}
$$

$$
f(x) = {}_{-k}C_x p^k (p-1)^x
$$

- 二項係数を負に拡張した場合の二項定理：

$$
\sum_{x=0}^{\infty} {}_{a}C_x t^x = (1+t)^a
$$

負の二項分布 (続き)

$$f(x) = {}_{-k}C_x p^k (p-1)^x$$

- $f(x)$ が確率関数であることの証明
  - $f(x) \geq 0$ は明らか
  - 二項定理

$$\sum_x {}_aC_x t^x = (1+t)^a$$

で $a = -k,\ t = p-1$ とおけば,

$$\sum_x {}_{-k}C_x (p-1)^x = p^{-k}$$

これより $\displaystyle\sum_x f(x) = p^k \sum_x {}_{-k}C_x (p-1)^x = 1$

負の二項分布の性質

- 期待値:

$$E(X) = \frac{k(1-p)}{p}$$

- 分散:

$$V(X) = \frac{k(1-p)}{p^2}$$

- 積率母関数

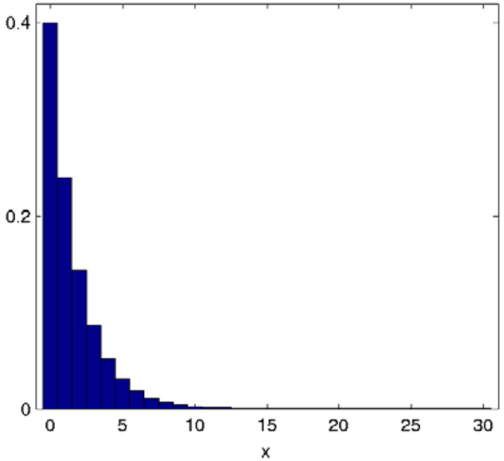
$$M_X(t) = p^k\{1-(1-p)e^t\}^{-k}$$

- 証明は宿題

負の二項分布の例

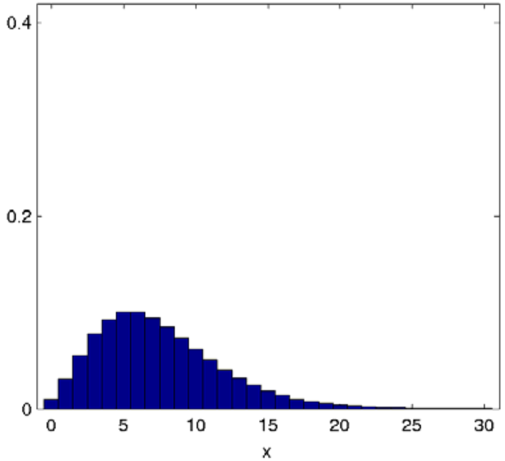
■ 成功回数を変えたとき

$$p = 0.4$$

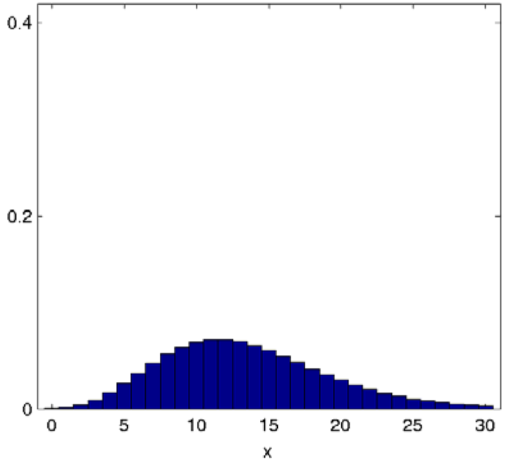


$$k = 1$$

成功回数少ない



$$k = 5$$



$$k = 9$$

成功回数多い

負の二項分布

- ある実験の成功と失敗の確率がそれぞれ p と $1-p$
- k 回目の成功を得るまでの失敗の回数を x で表せば，その確率は

$$f(x) = {}_{k+x-1}C_x p^k (1-p)^x$$

$$\text{for } x = 0, 1, 2, \ldots$$

- これを**負の二項分布**(negative binomial distribution)，または**パスカル分布**(Pascal distribution)という
- $k=1,\ y=x-1$ のとき，$f(y)$ は幾何分布と一致

幾何分布

$$f(x) = {}_{k+x-1}C_x p^k (1-p)^x \qquad \text{for } x = 0, 1, 2, \ldots$$

- 負の二項分布において $k=1,\ x \leftarrow x-1$ とおいたものを特に**幾何分布**(Geometric distribution)とよぶ

$$f(x) = p(1-p)^{x-1} \quad \text{for } x = 1, 2, \ldots$$

- ベルヌーイ試行の回数をあらかじめ決めないで，初めて成功するまでの試行回数 x の分布
- 名前の由来：幾何数列（等比数列）だから