

確率と統計(○)

「離散型確率分布の例(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/>

講義計画(シラバス)

94

- 確率と統計の基礎
- 確率変数, 確率分布
- 積率, 積率母関数
- 離散型の確率分布の例
- 連続型の確率分布の例
- 確率不等式, 擬似乱数
- 多次元の確率分布
- 大数の法則, 中心極限定理
- 統計的推定, 仮説検定

主な離散型の確率分布

95

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

成功確率の低い場合の二項分布 96

- 二項分布で、実験の成功率 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 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.
- **例**: 成功率が $p = 0.00003$, 実験回数が $n = 100000$ の二項分布の期待値は

$$E(X) = np = 100000 \times 0.00003 = 3$$

- このとき, $x = 0, 1, 2, 3, 4, 5$ くらいの生起確率はそれほど小さくなさそう?
- しかし, その確率の計算は大変!

$$f(5) = {}_{100000}C_5 (0.00003)^5 (0.99997)^{99995}$$

二項分布の極限

97

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

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

$(n \rightarrow \infty)$

- 証明:

$$\begin{aligned} & \lim_{n \rightarrow \infty} {}_nC_x \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \lim_{n \rightarrow \infty} \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \frac{n!}{(n-x)!n^x} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \end{aligned}$$

次頁へ続く

二項分布の極限(続き)

98

前頁からのつづき:

- $$\lim_{n \rightarrow \infty} \frac{n!}{(n-x)!n^x} = \lim_{n \rightarrow \infty} \frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-x+1}{n}$$
$$= \lim_{n \rightarrow \infty} 1 \times \frac{1 - \frac{1}{n}}{1} \times \cdots \times \frac{1 - \frac{x}{n} + \frac{1}{n}}{1} = 1$$

- 自然対数の底の定義 $e = \lim_{t \rightarrow 0} (1+t)^{\frac{1}{t}}$ で, $t = -\frac{\lambda}{n}$ と

おけば,

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

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

これらより,
$$\lim_{n \rightarrow \infty} {}_n C_x \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \longrightarrow \frac{e^{-\lambda} \lambda^x}{x!} \quad (\text{Q. E. D.})$$

ポアソン分布

■ ポアソン分布(Poisson distribution):

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

- $Po(\lambda)$ で表す事が多い
- $f(x)$ 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

$$e^{\lambda} = 1 + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \cdots = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

より,

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

ポアソン分布の性質

100

■ 期待値: $E(X) = \lambda$

■ 分散: $V(X) = \lambda$

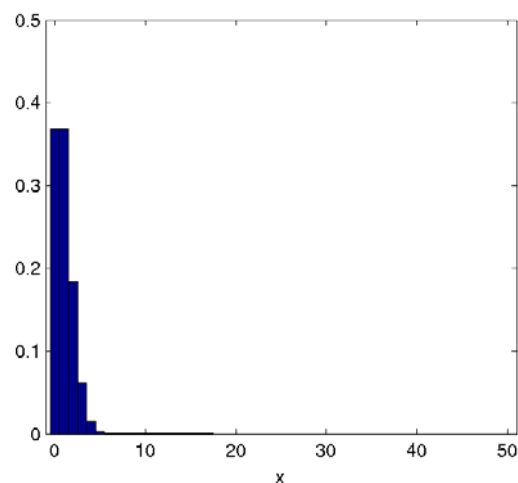
■ 積率母関数: $M_X(t) = \exp \{ \lambda(e^t - 1) \}$

■ 証明は演習

ポアソン分布の例

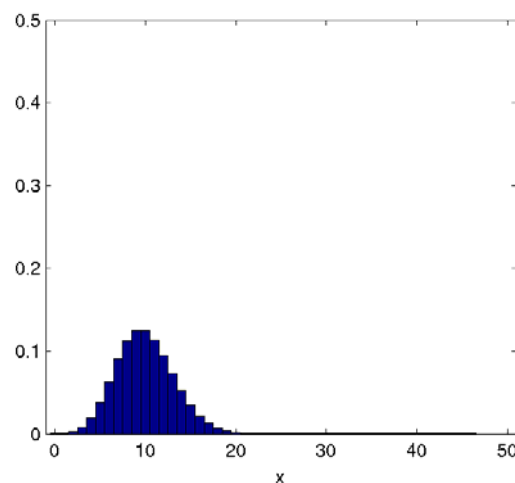
101

■ 実験の平均成功回数を変えたとき

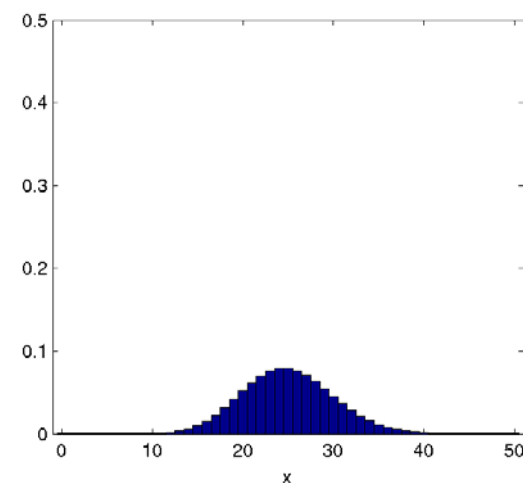


$$\lambda = 1$$

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(negative binomial distribution):
成功確率が p の実験に対して, k 回目の成功を得る
までの失敗の回数 x の確率分布

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

- $k + x$ 回目の試行は必ず成功.
 - それまでの $k + x - 1$ 回の試行中の失敗の
組み合わせ数は ${}_{k+x-1}C_x$.
- パスカル分布(Pascal distribution)とも呼ぶ.

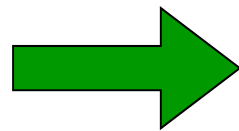
■ 負の二項分布の名前の由来: 負の二項係数

● 二項係数

$${}_a C_x = \frac{a \times (a-1) \times \cdots \times (a-x+1)}{x \times (x-1) \times \cdots \times 1} \quad a = 1, 2, \dots$$

● 負の二項係数

$$-_k C_x = (-1)^x {}_{k+x-1} C_x \quad k = 1, 2, \dots$$



$$f(x) = -_k C_x p^k (p-1)^x$$

■ 二項定理は負でも成り立つ:

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

負の二項分布(続き)

104

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

■ $f(x)$ 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

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

で $a = -k$, $t = p-1$ とおけば,

$$\sum_{x=0}^{\infty} {}_{-k}C_x (p-1)^x = p^{-k}$$

$$\text{これより } \sum_{x=0}^{\infty} f(x) = p^k \sum_{x=0}^{\infty} {}_{-k}C_x (p-1)^x = 1$$

負の二項分布の性質

105

■ 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:

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

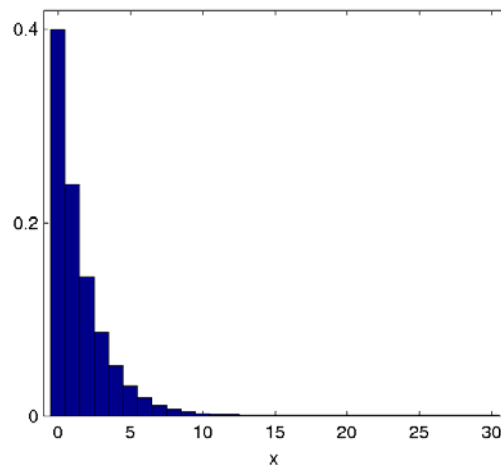
■ 証明は宿題

負の二項分布の例

106

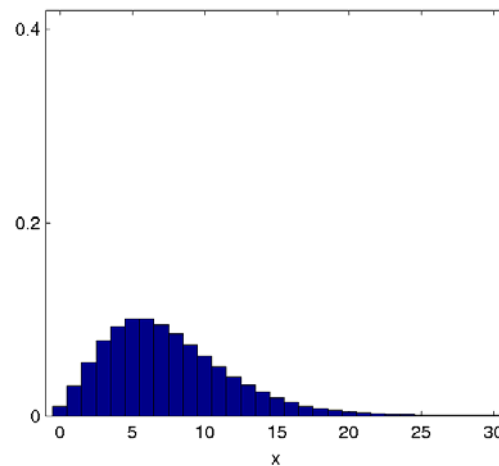
■ 成功回数を変えたとき

$$p = 0.4$$

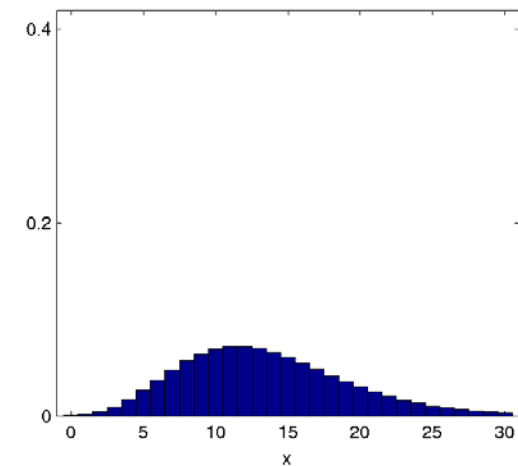


$$k = 1$$

成功回数少ない



$$k = 5$$



$$k = 9$$

成功回数多い

幾何分布

107

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

- 幾何分布(Geometric distribution): 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

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

- 負の二項分布において $k = 1$, $x \leftarrow x - 1$ とおいたものに対応
- 名前の由来: 幾何数列(等比数列)だから

- 離散分布
 - ポアソン分布
 - 負の二項分布
 - 幾何分布
- ポアソンの小数の法則
- 負の二項係数, 負の二項定理

- 負の二項分布の積率母関数, 期待値, 分散が次のように表せることを証明せよ.

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

$$E(X) = k(1 - p)/p$$

$$V(X) = k(1 - p)/p^2$$

ヒント:

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

$$E[X] = M'_X(0)$$

$$V(X) = M''_X(0) - M'_X(0)^2$$