統計データ解析 (I) 第 7 回

小池祐太

2017 年 11 月 8 日

中心極限定理

- $X_1, X_2, \ldots$ を独立同分布な確率変数列とする
  - データ解析の文脈では, X_i が「i 番目の観測データ」に対応
- n 番目までのデータの平均 (標本平均)

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

を使って X_1 の平均 μ (データの「真の」平均) を推測する問題を考える

- $X_1, X_2, \ldots$ は同分布なので, すべての確率変数が共通の平均 μ をもつことに注意

中心極限定理

- 前回説明した大数の法則は, サンプル数 n を大きくするに従い標本平均 $\bar{X}_n$ が真の平均 μ に限りなく近づくことを保証している
- 言い換えると,「推定誤差」$\bar{X}_n - \mu$ は n を大きくすると限りなく 0 に近づく
- しかし, 実際に推定誤差 $\bar{X}_n - \mu$ がどの程度の大きさになるのか定量的に評価する手段は与えていない

中心極限定理

- 統計学では, 推定誤差がある区間 $[\alpha, \beta]$ に入る確率

$$P(\alpha \leq \bar{X}_n - \mu \leq \beta) \tag{1}$$

  を計算することによって推定誤差を定量的に評価する (詳しい手順については「推定」の講義で説明する)
- 確率 (1) の正確な計算は一般には困難であるが, サンプル数が十分大きい場合ある関数の定積分で近似できることが知られている
- このことを具体的に述べたのが次の**中心極限定理**である

中心極限定理

定理 1 (中心極限定理)

$X_1, X_2, \ldots$ を独立同分布な確率変数列とし，その平均を μ，標準偏差を σ とする．このとき，すべての実数 $a < b$ に対して

$$P\left(a \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq b\right) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \qquad (n \to \infty)$$

が成り立つ．

中心極限定理

- 中心極限定理より，サンプル数 n が十分大きければ，確率

$$P\left(a\frac{\sigma}{\sqrt{n}} \leq \bar{X}_n - \mu \leq b\frac{\sigma}{\sqrt{n}}\right)$$

は定積分

$$\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx \tag{2}$$

によって近似できる

- 積分 (2) の被積分関数 $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ は**標準正規密度 (関数)** と呼ばれており，R では関数 `dnorm( )` で計算できる

中心極限定理

- 定積分 (2) 自体は関数 `pnorm( )` を用いてコマンド

```
pnorm(b)-pnorm(a)
```

  で計算できる
- $\boxed{\text{補足}}$
  - 詳細は次回以降の講義で説明するが, 各実数 $a \le b$ について区間 $[a, b]$ に値が入る確率が定積分 (2) で与えられるような確率変数を**標準正規確率変数**と呼び, 標準正規確率変数の値の分布の仕方を**標準正規分布**と呼ぶ
  - 中心極限定理の意味するところは, X_i たちの分布が何であっても, サンプル数 n が十分大きければ, 標本平均を正規化した量の $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ の分布の仕方は標準正規確率分布で近似できるということである

中心極限定理

- 中心極限定理のシミュレーションによる確認は, ヒストグラムによる可視化を用いる方法がよく利用される
- これは, $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ をシミュレーションした際のヒストグラムのビン $[a, b]$ における高さが

$$\frac{1}{b-a} P\left(a \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq b\right)$$

で与えられることを利用する (関数 `hist()` でオプション `freq` を `FALSE` に指定した場合)

中心極限定理

- 従ってビンの幅 $b-a$ が十分小さければ, 中心極限定理が正しい限りビン $[a,b]$ におけるヒストグラムの高さは $\phi(a)$ で近似できるはずである
- 従って, $\sqrt{n}(\bar{X}_n-\mu)/\sigma$ のヒストグラムに標準正規密度 $\phi(x)$ を重ね書きすることで, 近似の度合いを評価できる
- 実行例 `CLT.r`

中心極限定理

演習 1

中心極限定理の実行例 `CLT.r` において, 乱数の初期値・出現確率の変更・乱数の変更などを行っても中心極限定理が成立することを確認せよ. また, サンプル数の増減によって精度がどの程度変わるか観察せよ.

重複対数の法則

定理 2 (重複対数の法則)

$X_1, X_2, \dots$ を独立同分布な確率変数列とし，その平均を μ，標準偏差を σ とする．このとき，

$$\limsup_{n\to\infty} \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{2\sigma^2 \log\log n}} = 1 \quad \text{a.s.,} \tag{3}$$

$$\liminf_{n\to\infty} \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{2\sigma^2 \log\log n}} = -1 \quad \text{a.s.} \tag{4}$$

が成り立つ．より一般に，列

$$\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{2\sigma^2 \log\log n}} \right)_{n=3}^{\infty}$$

のある部分列の収束先となるような実数全体の集合を C とすると，C が閉区間 $[-1, 1]$ に一致する確率は 1 である．

重複対数の法則

- 定理 2 の後半の主張は **Hartman-Wintner の定理**として知られている
- 実行例 `LIL.r`

重複対数の法則

演習 2

重複対数の法則の実行例 `LIL.r` において，乱数の初期値・出現確率の変更・乱数の変更などを行っても重複対数の法則が成立することを確認せよ．また，サンプル数の増減によって精度がどの程度変わるか観察せよ．

少数の法則

- 少数の法則とは, 滅多に起こらない事象が起こる回数の分布に関する法則である
- 例えば，ある製品の不良品率 p はとても小さいとする．一日に n 個 (非常に多数とする) 生産するとき, 不良品は平均的には $\lambda = np$ 個発生するが, 日によって不良品の個数 S_n には多少のばらつきが生じる
- 従って S_n は確率変数であるが, S_n がとる値の確率法則は, 強度 λ の **Poisson 分布**で近似できることが知られている
- これを正確に述べたのが次の**少数の法則**である

少数の法則

定理 3 (少数の法則)

$X_1, X_2, \ldots, X_n$ を独立な確率変数列とし, 各 $i = 1, 2, \ldots, n$ について X_i は確率 $p_{n,i}$ で 1 を, 確率 $1 - p_{n,i}$ で 0 をとるとする:

$$P(X_i = 1) = p_{n,i}, \qquad P(X_i = 0) = 1 - p_{n,i} \qquad (i = 1, 2, \ldots, n).$$

このとき, ある正の実数 λ が存在して, $n \to \infty$ のとき

$$\max_{i=1,2,\ldots,n} p_{n,i} \to 0, \qquad \sum_{i=1}^{n} p_{n,i} \to \lambda$$

が成り立つならば, 任意の 0 以上の整数 k に対して

$$P\left(\sum_{i=1}^{n} X_i = k\right) \to e^{-\lambda} \frac{\lambda^k}{k!} \qquad (n \to \infty)$$

が成り立つ.

少数の法則

- 上の定理において，$\sum_{i=1}^{n} X_i$ が本小節冒頭の例の S_n に対応する
- $\boxed{\text{補足}}$
  - 詳細は次章で説明するが，取りうる値が 0 以上の整数全体で，値が整数 $k \geq 0$ となる確率が
    $$e^{-\lambda}\frac{\lambda^k}{k!} \tag{5}$$
    で与えられる確率変数を強度 λ の **Poisson 型確率変数**と呼び，その値の分布の仕方を強度 λ の Poisson 分布と呼ぶ
  - (5) は関数 `dpois()` で計算できる
- 実行例 `LRE.r`

少数の法則

演習 3

少数の法則の実行例 `LRE.r` において, 乱数の初期値を変更しても少数の法則が成立することを確認せよ. また, パラメータ n, p の変更によって結果がどのように変わるか観察せよ.

確率分布

- X を確率変数とする
  - 値がランダムに決定される変数で，すべての実数 $a \leq b$ に対して，その値が区間 $[a, b]$ に含まれる確率があらかじめ定められているような変数
- 各区間 $[a, b]$ $(a \leq b)$ と，X が区間 $[a, b]$ に含まれる確率

$$P(a \leq X \leq b)$$

との対応を示したものを，X の**確率分布**または単に**分布**といい，X はこの分布に**従う**という

確率分布

- 観測の結果として定まる確率変数の実現値はランダムに決定されるため, その値自体には格別の意味はなく, 現象の理解のためには値の出現しやすさの方にこそ興味がある
- そのため, 確率統計学では, 確率分布の数学的モデリングを通じて現象の理解を試みることとなる
- ここでは, いくつかの基本的な確率分布の数理モデルを, R におけるシミュレーション方法とあわせて説明する

離散分布

- 取りうる値が有限個, もしくは可算無限個 (例えば整数値のみとる場合) であるような確率変数は**離散型**であるといい, 対応する確率分布を**離散分布**と呼ぶ
- 離散分布は, その分布に従う確率変数 X が取りうる値 x のそれぞれに対して, $X = x$ となる確率 $P(X = x)$ を対応させる関数 $f(x) = P(X = x)$ を考えることで完全に決定される
- この関数 f を**確率質量関数**, あるいは単に**確率関数**と呼ぶ

離散分布

- 取りうる値が有限個の確率変数の場合と同様に, 離散型の確率変数に対してもその**平均**を定義したい
- いまの場合, 確率変数の取りうる値が無限個あるかもしれないため, その定義には少し注意を要する
- まず, 有限もしくは可算無限個の要素をもつ集合 $\mathcal{X}$ とその上で定義された実数値関数 φ に対して, 級数

$$\sum_{x\in\mathcal{X}} \varphi(x)$$

を定義することから始める[1]

[1] $\mathcal{X}$ は空集合でないとする

離散分布

- まず $\mathcal{X}$ が n 個の要素をもつ有限集合の場合, $\mathcal{X}$ の要素の適当な番号づけを $x_1, \ldots, x_n$ とし,

$$\sum_{x \in \mathcal{X}} \varphi(x) := \sum_{i=1}^{n} \varphi(x_i)$$

  と定義する
  - 加法の可換性より, この定義は $\mathcal{X}$ の要素の番号付けの仕方によらない

離散分布

- 次に, $\mathcal{X}$ が可算無限集合の場合, $\mathcal{X}$ の要素のある番号付け $x_1, x_2, \ldots$ に対して級数 $\sum_{i=1}^{\infty} \varphi(x_i)$ が絶対収束するとき,

$$\sum_{x \in \mathcal{X}} \varphi(x) := \sum_{i=1}^{\infty} \varphi(x_i)$$

  と定義し, 絶対収束しない場合は定義できないとする
  - 級数 $\sum_{i=1}^{\infty} \varphi(x_i)$ が絶対収束する場合, 正項級数の性質よりこの級数は $\mathcal{X}$ の要素の番号付けの仕方によらずに絶対収束し, 上式右辺の級数の値は番号付けの仕方によらずに定まる

離散分布

- 以上の準備の下, 離散型の確率変数 X の平均を以下のように定義する
- X の取りうる値全体からなる集合を $\mathcal{X}$ とする
- 級数 $\sum_{x\in\mathcal{X}} xP(X=x)$ が定義できるとき, X の**平均**を

$$E[X] := \sum_{x\in\mathcal{X}} xP(X=x)$$

で定義する
  - 平均は**期待値**とも呼ばれる
  - 級数 $\sum_{x\in\mathcal{X}} xP(X=x)$ が定義できない場合, X は平均をもたない

- より一般に, X の関数 $\varphi(X)$ に対して, 級数 $\sum_{x\in\mathcal{X}}\varphi(x)P(X=x)$ が定義できるとき, $\varphi(X)$ の期待値を

$$E[\varphi(X)] := \sum_{x\in\mathcal{X}}\varphi(x)P(X=x)$$

で定義する
- 特に, 正の整数 p に対して,

$$E[X^p] = \sum_{x\in\mathcal{X}} x^p P(X=x)$$

であり, これを p 次の**モーメント**あるいは**積率**と呼ぶ
  - 級数 $\sum_{x\in\mathcal{X}} x^p P(X=x)$ が定義できないとき, X は p 次のモーメントをもたない
  - 一般に, ある正整数 p に対して X が p 次のモーメントをもてば, $q \le p$ なるすべての正整数 q に対して X は q 次のモーメントをもつことが知られている

離散分布

- 前章と同様に, 離散型の確率変数 X に対して, 平均からのばらつき具合を定量化した指標として, X の**分散**を

$$\mathrm{Var}[X] := E[(X - E[X])^2] = \sum_{x \in \mathcal{X}} (x - E[X])^2 P(X = x)$$

で定義する (分散は X が 2 次のモーメントをもつときのみ定義できる)
- 分散の平方根 $\sqrt{\mathrm{Var}[X]}$ を**標準偏差**と呼ぶ

離散分布

- 離散分布の平均, モーメント, 分散, 標準偏差は, その分布に従う確率変数の平均, モーメント, 分散, 標準偏差で定義する
- 定義より明らかなように, 離散型の確率変数の平均, モーメント, 分散, 標準偏差はその分布のみに依存して定まるため, この定義は確率変数の選び方によらない
- むしろ, 確率変数の平均, モーメント, 分散, 標準偏差はその確率変数が従う分布のものとみなす方が本質的である

離散分布

- 前章では有限個の値をとる確率変数列のみについて大数の法則, 中心極限定理および重複対数の法則を説明したが, 実際にはそれらの主張は, 確率変数たちが 2 次のモーメントをもつ限り, 離散型の確率変数の列についても成り立つ
  - 2 次のモーメントをもたない場合, 中心極限定理と重複対数の法則は成立しない (そもそも分散が定義できない)
  - 大数の強法則は平均が存在すれば成立する
- ただし, 一般の場合の確率変数列の独立性と同分布性は以下のように定義する

離散分布

- まず, n 個の確率変数 $X_1, X_2, \ldots, X_n$ が**独立**であるとは, $a_i \leq b_i$ $(i = 1, \ldots, n)$ なる任意の実数 $a_1, b_1, \ldots, a_n, b_n$ に対して,

$$P(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2, \ldots, a_n \leq X_n \leq b_n) \\ = P(a_1 \leq X_1 \leq b_1)P(a_2 \leq X_2 \leq b_2) \cdots P(a_n \leq X_n \leq b_n) \quad (6)$$

  が成り立つことをいう
  - (6) の左辺は「X_1 が区間 $[a_1, b_1]$ に値をとり, X_2 が区間 $[a_2, b_2]$ に値をとり, $\ldots$, X_n が区間 $[a_n, b_n]$ に値をとる」という事象が起きる確率を表す

離散分布

- 次に, $X_1, X_2, \dots, X_n$ が**同分布**であるとは, $a_i \le b_i$ $(i = 1, \dots, n)$ なる任意の実数 $a_1, b_1, \dots, a_n, b_n$ に対して,

$$P(a_1 \le X_1 \le b_1) = P(a_2 \le X_2 \le b_2) = \cdots = P(a_n \le X_n \le b_n)$$

  が成り立つことをいう
- 確率変数の無限列に対する独立性および同分布性の定義は前章と同様
- 離散型の確率変数列の場合, これらの定義は $a_i = b_i$ $(i = 1, \dots, n)$ の場合のみ確認すれば十分であることがわかるため, 前章での定義と同じ形式となる

離散一様分布

- $x_1, \ldots, x_n$ を相異なる実数とする
- 取りうる値が $x_1, \ldots, x_n$ であり，確率関数が

$$f(x) = \frac{1}{n}, \qquad x \in \{x_1, \ldots, x_n\}$$

で与えられる離散分布を，集合 $\{x_1, \ldots, x_n\}$ 上の**離散一様分布**と呼ぶ

- 平均は $\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i$，分散は $\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ で与えられる

離散一様分布

実際, X を集合 $\{x_1, \ldots, x_n\}$ 上の離散一様分布に従う確率変数とすると,

$$E[X] = \sum_{i=1}^{n} x_i f(x_i) = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x},$$

$$\mathrm{Var}[X] = \sum_{i=1}^{n} (x_i - E[X])^2 f(x_i) = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

離散一様分布

- 例えば, 歪みのないサイコロを 1 回投げたときに出る目の分布は, 集合 $\{1, \ldots, 6\}$ 上の離散一様分布に従う
- 離散一様分布に従う乱数の発生には, 前章で説明した関数 `sample()` が利用できる (オプション `replace` を `TRUE` に指定して使う)
- 実行例 `sample2.r`

二項分布

- n を正の整数, p を 0 以上 1 以下の実数とする
- 取りうる値が $0, 1, \ldots, n$ であり, 確率関数が

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, \ldots, n$$

で与えられる離散分布を, 試行回数 n, 成功確率 p の**二項分布**と呼ぶ
- 平均は np, 分散は $p(1-p)$ で与えられる

実際, X を試行回数 n, 成功確率 p の二項分布に従う確率変数とすると,

$$
\begin{aligned}
E[X] &= \sum_{x=0}^{n} x f(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n-1} \frac{n!}{x!(n-x-1)!} p^{x+1} (1-p)^{n-x-1} \\
&= np \sum_{x=0}^{n-1} \binom{n-1}{x} p^x (1-p)^{n-1-x} \\
&= np
\end{aligned}
$$

であり，また，

$$
\begin{aligned}
E[X^2] &= \sum_{x=0}^{n} x^2 f(x) = \sum_{x=0}^{n} x(x-1)\binom{n}{x} p^x (1-p)^{n-x} + E[X] \\
&= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} + np \\
&= \sum_{x=0}^{n-2} \frac{n!}{x!(n-x-2)!} p^{x+2} (1-p)^{n-x-2} + np \\
&= n(n-1)p^2 \sum_{x=0}^{n-2} \binom{n-2}{x} p^x (1-p)^{n-2-x} + np \\
&= n(n-1)p^2 + np = (np)^2 + np(1-p)
\end{aligned}
$$

となるから，

$$
\mathrm{Var}[X] = E[X^2] - (E[X])^2 = (np)^2 + np(1-p) - (np)^2 = np(1-p)
$$

二項分布

- 特に, 試行回数 1 の二項分布を **Bernoulli 分布**と呼ぶ
- 例えば, 表が出る確率が p のコインを n 回投げたときに表が出る回数は試行回数 n, 成功確率 p の二項分布に従う
- 前回も述べたように, 二項分布に従う乱数の発生には関数 `rbinom()` を用いる
- なお, 原則として, ある確率分布に従う乱数を生成するための R の関数の命名規則は, 「`r` + その乱数が従う分布の名前の省略形」となっている (離散一様分布など一部例外がある)
- また, 離散分布の場合, その確率関数を計算するための関数が, 同じ省略形の文頭に `d` をつけることで得られる
  - 例えば, 二項分布の確率関数は関数 `dbinom()` で計算できる
- 実行例 `rbinom2.r`

二項分布

演習 4

$X_1, \ldots, X_n$ を成功確率 p の Bernoulli 分布に従う独立同分布な確率変数列とすると, $\sum_{i=1}^n X_i$ の分布は試行回数 n, 成功確率 p の二項分布に従うことが知られている．このことを上の実行例を参考にしながらグラフを描画して確認せよ．

Poisson 分布

- λ を正の実数とする
- 取りうる値が 0 以上の整数であり, 確率関数が

$$f(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \qquad x = 0, 1, \ldots$$

  で与えられる離散分布をパラメーター λ の **Poisson 分布**と呼び, 記号 $P_o(\lambda)$ で表す
  - λ は**強度**と呼ばれることがある
- 平均, 分散はともに λ で与えられる

実際, X をパラメーター λ の Poisson 分布に従う確率変数とすると,

$$E[X] = \sum_{x=0}^{\infty} xf(x) = \sum_{x=0}^{\infty} x\frac{\lambda^x}{x!}e^{-\lambda} = e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!}$$
$$= \lambda e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = \lambda$$

であり, また,

$$E[X^2] = \sum_{x=0}^{\infty} x^2 f(x) = \sum_{x=0}^{\infty} x(x-1)\frac{\lambda^x}{x!}e^{-\lambda} + E[X]$$
$$= e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^x}{(x-2)!} + \lambda = \lambda^2 e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} + \lambda = \lambda^2 + \lambda$$

となるから,

$$\mathrm{Var}[X] = E[X^2] - (E[X])^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Poisson 分布

- 放射性物質から一定時間に放射される粒子の数や，一定期間に起こる交通事故の数などは Poisson 分布に従うことが知られている
- また，上で観察したように，発生確率が低い事象が十分長い期間のあいだに起こる回数の分布は Poisson 分布で近似できる
- Poisson 分布に従う乱数の発生には関数 `rpois()` を用いる
- 実行例 `rpois2.r`

Poisson 分布

演習 5

X, Y を独立な 2 つの確率変数とし，それぞれパラメーター λ_1, λ_2 の Poisson 分布に従うとする．このとき，和 $X + Y$ の分布はパラメーター $\lambda_1 + \lambda_2$ の Poisson 分布に従うことが知られている．このことを上の実行例を参考にしながらグラフを描画して確認せよ．

幾何分布

- $0 < p \leq 1$ とする
- 取りうる値が 0 以上の整数であり, 確率関数が

$$f(x) = p(1-p)^x, \qquad x = 0, 1, \ldots$$

で与えられる離散分布を成功確率 p の**幾何分布**と呼ぶ
- 平均は $(1-p)/p$, 分散は $(1-p)/p^2$ で与えられる

実際, X を成功確率 p の幾何分布に従う確率変数とすると,

$$
\begin{aligned}
E[X] &= \sum_{x=0}^{\infty} x f(x) = \sum_{x=0}^{\infty} x p(1-p)^x = p(1-p)\sum_{x=1}^{\infty} x(1-p)^{x-1} \\
&= -p(1-p)\frac{d}{dp}\sum_{x=1}^{\infty}(1-p)^x = -p(1-p)\frac{d}{dp}\frac{1-p}{p} \\
&= -p(1-p)\cdot\left(-\frac{1}{p^2}\right) = \frac{1-p}{p}
\end{aligned}
$$

であり，また，

$$
\begin{aligned}
E[X^2] = \sum_{x=0}^{\infty} x^2 f(x) &= \sum_{x=0}^{\infty} x(x-1)p(1-p)^x + E[X] \\
&= p(1-p)^2 \sum_{x=2}^{\infty} x(x-1)p(1-p)^{x-2} + \frac{1-p}{p} \\
&= p(1-p)^2 \frac{d^2}{dp^2} \sum_{x=2}^{\infty} (1-p)^x + \frac{1-p}{p} \\
&= p(1-p)^2 \frac{d^2}{dp^2} \frac{(1-p)^2}{p} + \frac{1-p}{p} \\
&= p(1-p)^2 \frac{2}{p^3} + \frac{1-p}{p} = 2\left(\frac{1-p}{p}\right)^2 + \frac{1-p}{p}
\end{aligned}
$$

となるから，

$$
\mathrm{Var}[X] = E[X^2] - (E[X])^2 = \left(\frac{1-p}{p}\right)^2 + \frac{1-p}{p} = \frac{1-p}{p^2}
$$

幾何分布

- 表が出る確率が p のコインを投げ続けて，初めて表が出るまでに出た裏の回数は，成功確率 p の幾何分布に従う
- 幾何分布に従う乱数の発生には関数 `rgeom()` を用いる
- 実行例 `rgeom.r`