

クレジット:

UTokyo Online Education 統計データ解析 I 2017 小池祐太

ライセンス:

利用者は、本講義資料を、教育的な目的に限ってページ単位で利用することができます。特に記載のない限り、本講義資料はページ単位でクリエイティブ・コモンズ 表示-非営利-改変禁止 ライセンスの下に提供されています。

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

本講義資料内には、東京大学が第三者より許諾を得て利用している画像等や、各種ライセンスによって提供されている画像等が含まれています。個々の画像等を本講義資料から切り離して利用することはできません。個々の画像等の利用については、それぞれの権利者の定めるところに従ってください。



統計データ解析 I (平成 29 年度)

東京大学大学院数理科学研究科
統計データ解析教育研究グループ

村田 昇 (早稲田大学, 東京大学)

吉田朋広 (東京大学)

小池祐太 (首都大学東京, 東京大学)

第8章 推定

分析対象の観測データに基づいて統計解析を行う際、確率統計学では観測データのある確率変数たちの実現値と考えてモデル化する。このとき、確率変数たちの従う分布のもつなんらかの特性量(平均や分散など)を評価したり、分布そのものを決定することが統計解析の目標の1つとなるが、この作業を一般に**推定**と呼ぶ。本章の目的は、統計学で広く利用されている代表的な推定方法を説明することである。

以下本章では、観測データが独立同分布な確率変数列 X_1, X_2, \dots, X_n がモデル化されている状況を考える。この場合、 X_i たちが従う共通の分布 \mathcal{L} に関する推定を行うことが目標となるが、この分布としてすべての分布を考察対象とすると、対象とする範囲が広くなりすぎて、サンプル数 n が十分大きくない限り応用上意味のある結論を導き出すことが困難になる。そのため、以下では確率分布 \mathcal{L} を特徴づけるなんらかのパラメーター θ を考察対象とする(例えば \mathcal{L} の平均、分散、歪度、尖度など)。

8.1. 点推定

\mathcal{L} に含まれるあるパラメーター θ を X_1, \dots, X_n のある関数

$$\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$$

で推定することを、**点推定**と呼ぶ。このとき、 $\hat{\theta}$ は θ の**推定量**と呼ばれる。点推定については、すでに前章で記述統計量との関連で議論した。そこでみたように、多くの記述統計量は、適当なモデル化の下でなんらかのパラメーターの推定量とみなせる。例えば、 \mathcal{L} の平均 μ を標本平均 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ によって推定することが点推定であり、 \bar{X} は μ の推定量となる。

一般に、1つのパラメーターに対して推定量は無数に存在するため、使うべき推定量を決定するために、推定量の良さを評価する基準が必要である。前章で議論したように、そのような基準として代表的なものに、**不偏性**と**一致性**がある： $\hat{\theta}$ が θ の不偏推定量であるとは、 $\hat{\theta}$ の平均が θ 、すなわち

$$E[\hat{\theta}] = \theta$$

が成り立つことをいい、 $\hat{\theta}$ が θ の(強)一致推定量であるとは、 $n \rightarrow \infty$ のとき $\hat{\theta}$ が θ に収束する確率が1であることをいう。

シンプルな状況では、1つのパラメーターに対して複数の不偏推定量が存在する場合も起こりうる。例えば、 \mathcal{L} の平均 μ の不偏推定量としては、標本平均 \bar{X} 以外にも、例えば X_1 が考えられる。より“自然な”推定量の例を挙げると、 \mathcal{L} が直線 $x = \mu$ に関して線対称な密度をもつ連続分布であったならば、 X_1, \dots, X_n のメディアンも μ の不偏推定量である。このような場合、不偏推定量の中から使うべきものを決定するための基準を設ける必要がある。自然な基準として、推定値のばらつき(分散)が最も小さいものを選ぶという基準が考えられる。すなわち、パラメーター θ の不偏推定量 $\hat{\theta}$ で、 θ の任意の不偏推定量 $\hat{\theta}'$ に対して

$$\text{Var}[\hat{\theta}] \leq \text{Var}[\hat{\theta}']$$

を満たすようなものを選ぶということである。このような推定量 $\hat{\theta}$ を**一様最小分散不偏推定量**と呼ぶ。一様最小分散不偏推定量を見出す方法の1つとして、次の結果が利用できる： \mathcal{L} は1つの(1次元)パラメーター θ を含む連続分布であるとし、その確率

密度関数 $f_\theta(x)$ は θ に関して偏微分可能であるとする。このとき、緩やかな仮定の下で、 θ の任意の不偏推定量 $\hat{\theta}$ に対して以下の不等式が成り立つ:

$$(8.1) \quad \text{Var}[\hat{\theta}] \geq \frac{1}{nI(\theta)}.$$

ただし,

$$I(\theta) = \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial \theta} \log f_\theta(x) \right)^2 f_\theta(x) dx.$$

不等式 (8.1) は **Cramér-Rao の不等式** と呼ばれ、その下界 $1/(nI(\theta))$ は **Cramér-Rao 下界** と呼ばれる。また、 $I(\theta)$ は **Fisher 情報量** と呼ばれる。Cramér-Rao の不等式より、もし θ の不偏推定量 $\hat{\theta}$ で分散が Cramér-Rao 下界 $1/(nI(\theta))$ に一致するものが存在すれば、それは一様最小分散不偏推定量となる。

例 8.1 (正規分布の平均の推定). \mathcal{L} が平均 μ 、分散 σ^2 の正規分布でモデル化されているとする。このとき、平均パラメーター μ に関する Fisher 情報量は

$$I(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sigma^4} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sigma^2}$$

となるから、Cramér-Rao 下界は σ^2/n となる。従って、標本平均 \bar{X} の分散は Cramér-Rao 下界に一致するので、 \bar{X} は μ の一様最小分散不偏推定量である。

```
> ## 標本平均が一様最小分散不偏推定量であることの確認
> set.seed(123) # 乱数の初期値を指定
> mu <- 10
> sigma <- 5
> mean.est <- function(n, mu, sigma){
+   x <- rnorm(n, mean = mu, sd = sigma)
+   xbar <- mean(x) # 標本平均
+   x1 <- x[1] # X1
+   med <- median(x) # メディアン
+   return(c(xbar, x1, med))
+ }
> n <- 10 # サンプル数
> MC <- 30000 # 実験回数
> mu.hats <- replicate(MC, mean.est(n, mu, sigma)) # mean.est(n) を MC 回実行して結果を記録
> mean(mu.hats[1, ]) # 標本平均の平均
[1] 10.00933
> mean(mu.hats[2, ]) # X1 の平均
[1] 10.01986
> mean(mu.hats[3, ]) # メディアンの平均
[1] 10.0121
> sigma^2/n # Cramer-Rao 下界
[1] 2.5
> var(mu.hats[1, ]) # 標本平均の分散
[1] 2.502021
> var(mu.hats[2, ]) # X1 の分散
[1] 25.15749
> var(mu.hats[3, ]) # メディアンの分散
[1] 3.457196
```

(umvue.r)

8.2. 最尤法

興味あるパラメーターが、平均や分散といった記述統計量と自然に関連づけられるパラメーターではない場合、推定量の構成が自明ではないことがある。このような場

合でも、確率分布 \mathcal{L} に対するモデルがいくつかのパラメータ $\theta_1, \theta_2, \dots, \theta_p$ を除いて特定されている状況であれば、一般的に適用可能な $\theta_1, \theta_2, \dots, \theta_p$ の推定量の構成方法がいくつか知られている。ここでは代表的な方法として最尤法を説明する。

まず \mathcal{L} が離散分布の場合を考え、その確率関数を $f_{\theta}(x)$ と書くことにする (確率関数のパラメータ $\theta := (\theta_1, \dots, \theta_p)$ への依存を明示するために添え字 θ をつけている)。このとき、パラメータ θ を一つ定めれば、観測値として $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ が得られる理論上の確率を

$$\prod_{i=1}^n f_{\theta}(x_i) = f_{\theta}(x_1) \cdot f_{\theta}(x_2) \cdots f_{\theta}(x_n)$$

で求めることができる。実際に得られている観測データは X_1, X_2, \dots, X_n であるから、パラメータ θ に対してそのような観測データが得られる理論上の確率は、

$$L(\theta) := \prod_{i=1}^n f_{\theta}(X_i)$$

で与えられる。 $L(\theta)$ は観測データ X_1, X_2, \dots, X_n が現れるのにパラメータ θ の値がどの程度尤もらしいか測る尺度と解釈でき、 θ の **尤度** と呼ばれる。 $L(\theta)$ を θ の関数とみなしたものを **尤度関数** と呼ぶ。 **最尤法** は観測データに対して「最も尤もらしい」パラメータを θ の推定量として採用する方法である。すなわち、尤度関数 $L(\theta)$ を最大化するパラメータ値 $\hat{\theta}$ を θ の推定量とする:

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta).$$

ここで、 Θ は尤度関数の定義域である。上の式は以下のようにも書かれる:

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L(\theta).$$

$\hat{\theta}$ は **最尤推定量** と呼ばれる。

なお、尤度関数は積の形をしていて扱いにくいので、和の形に直すために対数を取ることが多い:

$$\ell(\theta) := \log L(\theta) = \sum_{i=1}^n \log f_{\theta}(X_i).$$

対数関数は狭義増加であるから、 $\ell(\theta)$ の最大化と $L(\theta)$ の最大化は同義である。 $\ell(\theta)$ は **対数尤度関数** と呼ばれる。

例 8.2 (Poisson 分布の最尤推定). \mathcal{L} がパラメータ $\lambda > 0$ の Poisson 分布としてモデル化されている場合を考える。このとき、未知パラメータは λ であり、対数尤度関数は

$$\ell(\lambda) = \sum_{i=1}^n \log \frac{\lambda^{X_i}}{X_i!} e^{-\lambda} = \sum_{i=1}^n (X_i \log \lambda - \log X_i!) - n\lambda$$

で与えられる。いま、少なくとも 1 つの i について $X_i > 0$ であると仮定する。このとき、 $\ell(\lambda)$ を微分すると、

$$\ell'(\lambda) = \frac{1}{\lambda} \sum_{i=1}^n X_i - n, \quad \ell''(\lambda) = -\frac{1}{\lambda^2} \sum_{i=1}^n X_i < 0$$

を得る。従って方程式 $\ell'(\lambda) = 0$ の解が $\ell(\lambda)$ を最大化するから、 $\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i$ が λ の最尤推定量である。

\mathcal{L} が連続分布の場合は、確率関数の代わりに確率密度関数を用いて尤度を計算する。

例 8.3 (指数分布の最尤推定). \mathcal{L} がパラメータ $\lambda > 0$ の指数分布としてモデル化されている場合を考える. このとき, 未知パラメータは λ であり, 対数尤度関数は

$$\ell(\lambda) = \sum_{i=1}^n \log \lambda e^{-\lambda X_i} = n \log \lambda - \lambda \sum_{i=1}^n X_i$$

で与えられる. $\ell(\lambda)$ を微分すると,

$$\ell'(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^n X_i, \quad \ell''(\lambda) = -\frac{n}{\lambda^2} < 0$$

を得る. 従って方程式 $\ell'(\lambda) = 0$ の解が $\ell(\lambda)$ を最大化するから, $\hat{\lambda} = (\frac{1}{n} \sum_{i=1}^n X_i)^{-1}$ が λ の最尤推定量である.

例 8.4 (ガンマ分布の最尤推定). \mathcal{L} がパラメータ $\nu, \alpha > 0$ のガンマ分布としてモデル化されている場合を考える. このとき, 未知パラメータは ν, α であり, 対数尤度関数は

$$\begin{aligned} \ell(\nu, \alpha) &= \sum_{i=1}^n \log \frac{\alpha^\nu}{\Gamma(\nu)} X_i^{\nu-1} e^{-\alpha X_i} \\ &= n\nu \log \alpha - n \log \Gamma(\nu) + \sum_{i=1}^n \{(\nu - 1) \log X_i - \alpha X_i\} \end{aligned}$$

で与えられる. $\ell(\nu, \alpha)$ を最大化するような ν, α は解析的には求まらないため, 実際の計算では数値的に求めることになる (以下の実行例を参照). 数値計算の際に対数尤度関数の勾配 (偏導関数からなるベクトル) があると便利なので計算しておく:

$$\frac{\partial \ell}{\partial \nu}(\nu, \alpha) = n \log \alpha - n\psi(\nu) + \sum_{i=1}^n \log X_i, \quad \frac{\partial \ell}{\partial \alpha}(\nu, \alpha) = \frac{n\nu}{\alpha} - \sum_{i=1}^n X_i$$

ここに, $\psi(\nu) = \frac{d}{d\nu} \log \Gamma(\nu)$ であり, デイガンマ関数と呼ばれる (R では関数 `digamma()` で計算できる).

広い範囲の確率分布に対して, 最尤推定量は一致性を持つことが知られている.

```
> ### ガンマ分布の最尤推定
> ## 観測データから最尤推定量を計算する関数の作成
> ## x: 観測データ, nu0: nu の初期値, alpha0: alpha の初期値
> ## 数値的最適化のためには尤度関数を最初に評価する初期値が必要
> mle.gamma <- function(x, nu0 = 1, alpha0 = 1){
+
+   n <- length(x)
+
+   # 対数尤度関数のマイナスを計算する関数 (最小化を考えるため)
+   # suppressWarnings で定義域外で評価された際の警告を表示させないようにしている
+   f <- function(theta)
+     suppressWarnings(-sum(log(dgamma(x, theta[1], theta[2]))))
+
+   # f の勾配 (偏微分からなるベクトル) を計算する関数
+   # 指定しなくても大丈夫だが, 指定した方が計算が速いことが多い
+   gr <- function(theta){
+
+     f.nu <- -n*log(theta[2])+n*digamma(theta[1])-sum(log(x)) # nu に関する偏微分
+     f.alpha <- -n*theta[1]/theta[2]+sum(x) # alpha に関する偏微分
+
+     return(c(f.nu, f.alpha))
+   }
+
+   # 関数 f の最小化. "BFGS"は準 Newton 法の使用を指定
```

```

+   opt <- optim(c(nu = nu0, alpha = alpha0), fn = f, gr = gr,
+               method = "BFGS")
+
+   return(opt$par)
+ }
> ## シミュレーションによる一致性の検証
> set.seed(123) # 乱数の初期化
> nu <- 5
> alpha <- 2
> x <- rgamma(1000, shape = nu, rate = alpha)
> mle.gamma(x)

      nu      alpha
5.018777 2.061620

> ## 複数回行なった場合のヒストグラムの描画
> MC <- 1000 # 各実験の繰り返し回数
> nsample <- c(10, 100, 1000) # 実験するサンプル数
> op <- par(mfcol = c(2, 3)) # 画面を 2x3 に分割
> for(n in nsample){
+   thetihat <- replicate(MC, mle.gamma(rgamma(n, nu, alpha))) # 実験を MC 回実行
+   hist(thetihat[1, ], breaks = 20, col = "lightblue", border = "lightblue",
+        xlim = c(0, 20), main = paste0("n=",n), xlab = expression(nu))
+   abline(h = 0, lwd = 2, lty = "dotted")
+   abline(v = nu, col = "red", lwd = 2, lty = "dotted")
+   hist(thetihat[2, ], breaks = 20, col = "lightgreen", border = "lightgreen",
+        xlim = c(0, 10), main = paste0("n=",n), xlab = expression(alpha))
+   abline(h = 0, lwd = 2, lty = "dotted")
+   abline(v = alpha, col = "red", lwd = 2, lty = "dotted")
+ }
> par(op) # プロット環境の初期化

```

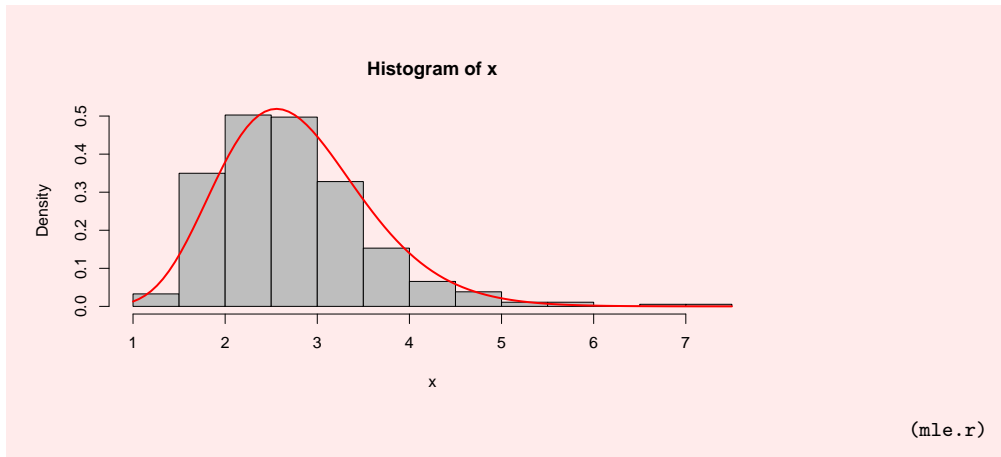
```

> ## 実データへの適用
> kikou <- read.csv("kikou2016.csv", fileEncoding = "sjis") # データの読み込み
> x <- kikou$風速 # 風速のデータにガンマ分布をあてはめてみる
> hist(x, freq = FALSE, breaks = 20, col = "gray") # ヒストグラムの描画
> (theta <- mle.gamma(x)) # 最尤推定

      nu      alpha
12.233964 4.392418

> curve(dgamma(x, theta[1], theta[2]), add = TRUE, col = "red",
+       lwd = 2) # あてはめたガンマ分布の密度を重ね書き

```



演習 8.1. 最尤法について調べてみよう.

- (1) 確率分布 \mathcal{L} が成功確率 p の幾何分布でモデル化される場合の, パラメーター p に対する最尤推定量を求めよ.
- (2) 確率分布 \mathcal{L} が平均 μ , 分散 σ^2 の正規分布でモデル化される場合の, パラメーター μ, σ^2 に対する最尤推定量を求めよ.

8.3. 区間推定

未知パラメーター θ を推定量 $\hat{\theta}$ で点推定した場合, 通常推定値は真のパラメーター値とは異なるため, 推定誤差が必ず存在する. そのため, 推定結果の定量的な評価には, 推定誤差の評価が重要となる. 統計学では, ある値 $\alpha \in (0, 1)$ を固定したとき,

$$P(l \leq \hat{\theta} - \theta \leq u) \geq 1 - \alpha$$

が成り立つような l, u を観測データから推定することで推定誤差の評価を試みる. 上の式の意味するところは, 「誤差 $\hat{\theta} - \theta$ が区間 $[l, u]$ の外側にある確率が α 以下」ということである. 上の式を変形すると,

$$P(\hat{\theta} - u \leq \theta \leq \hat{\theta} - l) \geq 1 - \alpha$$

となる. 従って, ここで行っているのは, パラメーター θ が含まれているような確率が $1 - \alpha$ 以上となるような区間 $[\hat{\theta} - u, \hat{\theta} - l]$ を推定することだと言い換えられる. このように, 未知パラメーターが含まれている確率があらかじめ決められたある値以上となるような区間を推定することを **区間推定** と呼ぶ.

より一般には, 未知パラメーター θ とある値 $\alpha \in (0, 1)$ に対して,

$$P(L \leq \theta \leq U) \geq 1 - \alpha$$

が成り立つような確率変数 L, U を観測データから求めることになる. このとき, 区間 $[L, U]$ を $100(1 - \alpha)\%$ **信頼区間**, L を $100(1 - \alpha)\%$ **下側信頼限界**, U を $100(1 - \alpha)\%$ **上側信頼限界**, $1 - \alpha$ を **信頼係数** とそれぞれ呼ぶ. 慣習として $\alpha = 0.01, 0.05, 0.1$ とすることが多い.

信頼区間は幅が狭いほど真のパラメーターが取りうる値の範囲を限定することになるため, 推定精度が良いといえる. 一方で, 信頼区間 $[L, U]$ の幅が狭いほど確率 $P(L \leq \theta \leq U)$ は小さくなるため, 最も推定精度の良い $100(1 - \alpha)\%$ 信頼区間 $[L, U]$ は,

$$P(L \leq \theta \leq U) = 1 - \alpha$$

を満たす. そのため, 実行可能である限り, $100(1 - \alpha)\%$ 信頼区間 $[L, U]$ の構成では上の式を満たすように L, U を決定する.

8.4. 正規母集団の場合の区間推定

この節では、 X_i たちの従う分布が平均 μ 、分散 σ^2 の正規分布の場合に、 μ および分散 σ^2 の区間推定をする方法を説明する。

8.4.1. 平均の区間推定 (分散既知). はじめに、分散 σ^2 がすでにわかっている場合に平均 μ の区間推定をする方法を説明する。これには次の結果を用いる:

命題 8.1. Z_1, Z_2, \dots, Z_k を独立な確率変数列とし、各 $i = 1, 2, \dots, k$ に対して Z_i は平均 μ_i 、分散 σ_i^2 の正規分布に従うとする。このとき、 a_0 を実数、 a_1, \dots, a_k を k 個の 0 でない実数とすると、 $a_0 + \sum_{i=1}^k a_i Z_i$ は平均 $a_0 + \sum_{i=1}^k a_i \mu_i$ 、分散 $\sum_{i=1}^k a_i^2 \sigma_i^2$ の正規分布に従う。

上の命題を、

$$k = n, \mu_i = \mu, \sigma_i^2 = \sigma^2, a_0 = 0, a_i = 1/n \quad (i = 1, \dots, n)$$

として適用すると、標本平均 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ は平均 μ 、分散 σ^2/n の正規分布に従うことがわかる。再び命題 8.1 を $k = 1, \mu_1 = \mu, \sigma_1^2 = \sigma^2/n, a_0 = -\sqrt{n}\mu/\sigma, a_1 = \sqrt{n}/\sigma$ として適用すると、確率変数

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$$

が標準正規分布に従うことがわかる。従って、 $\alpha \in (0, 1)$ を定めたとき、 $z_{1-\alpha/2}$ を標準正規分布の $100(1 - \alpha/2)\%$ 分位点とすれば、

$$P\left(-z_{1-\alpha/2} \leq \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq z_{1-\alpha/2}\right) = 1 - \alpha$$

が成り立つことがわかる。実際、 ϕ を標準正規分布の確率密度関数とすると、

$$\begin{aligned} P\left(-z_{1-\alpha/2} \leq \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq z_{1-\alpha/2}\right) &= \int_{-z_{1-\alpha/2}}^{z_{1-\alpha/2}} \phi(x) dx \\ &= \int_{-\infty}^{z_{1-\alpha/2}} \phi(x) dx - \int_{-\infty}^{-z_{1-\alpha/2}} \phi(x) dx \end{aligned}$$

となるが、ここで定義より

$$\int_{-\infty}^{z_{1-\alpha/2}} \phi(x) dx = 1 - \alpha/2$$

であり、また

$$\begin{aligned} \int_{-\infty}^{-z_{1-\alpha/2}} \phi(x) dx &= - \int_{\infty}^{z_{1-\alpha/2}} \phi(-y) dy \quad (y = -x \text{ と置換}) \\ &= \int_{z_{1-\alpha/2}}^{\infty} \phi(y) dy \quad (\phi \text{ は偶関数}) \\ &= \int_{-\infty}^{\infty} \phi(y) dy - \int_{-\infty}^{z_{1-\alpha/2}} \phi(y) dy \\ &= 1 - (1 - \alpha/2) = \alpha/2 \end{aligned}$$

(全確率 $\int_{-\infty}^{\infty} \phi(y) dy = 1$ に注意) であるから、

$$P\left(-z_{1-\alpha/2} \leq \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq z_{1-\alpha/2}\right) = 1 - \alpha/2 - \alpha/2 = 1 - \alpha$$

となる。カッコ内を μ について解くと、

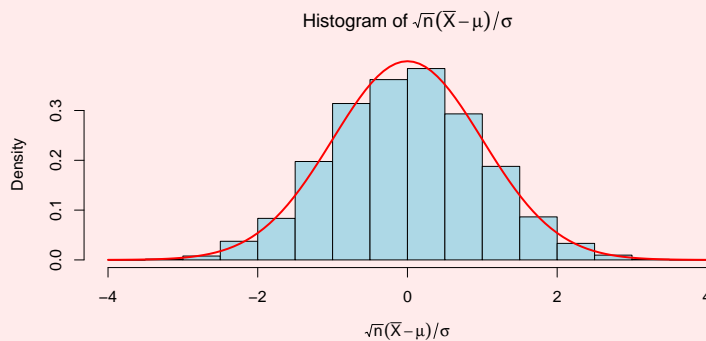
$$P(\bar{X} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n} \leq \mu \leq \bar{X} + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}) = 1 - \alpha$$

が得られるので、 σ が既知であれば、

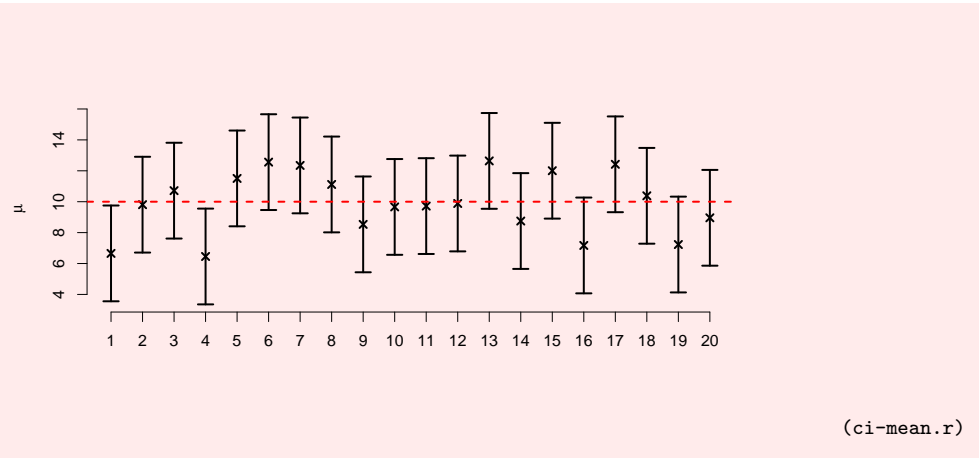
$$[\bar{X} - z_{1-\alpha/2} \cdot \sigma / \sqrt{n}, \bar{X} + z_{1-\alpha/2} \cdot \sigma / \sqrt{n}]$$

が平均 μ の $100(1 - \alpha)\%$ 信頼区間を与える。

```
> set.seed(111) # 乱数の初期値の設定
> mu <- 10 # 平均
> sigma <- 5 # 標準偏差
> # n 個の正規標本から標本平均と平均の 100(1-alpha)%信頼区間の上下端・下端を
> # 計算する関数の作成
> myest <- function(n, alpha = 0.05){
+
+   x <- rnorm(n, mean = mu, sd = sigma) # 正規乱数の生成
+   xbar <- mean(x) # 標本平均
+   ui <- xbar + qnorm(1-alpha/2) * sigma/sqrt(n) # 上側信頼限界
+   li <- xbar - qnorm(1-alpha/2) * sigma/sqrt(n) # 下側信頼限界
+
+   return(c(xbar = xbar, ui = ui, li = li))
+ }
> result <- replicate(10000, myest(10)) # n=10として10000回実験
> ## sqrt(n)*(xbar-mu)/sigmaが標準正規分布に従うことの確認
> hist(sqrt(10) * (result["xbar", ] - mu)/sigma, freq = FALSE,
+       col = "lightblue", xlab = expression(sqrt(n)*(bar(X)-mu)/sigma),
+       main = expression(paste("Histogram of ", sqrt(n)*(bar(X)-mu)/sigma)))
> curve(dnorm, add = TRUE, col = "red", lwd = 2)
```



```
> ## 真の平均 mu が 95%信頼区間に含まれている割合が約 95%であることの確認
> length(which(mu <= result["ui", ] & mu >= result["li", ]))/10000
[1] 0.9502
> ## はじめの 20 個の信頼区間の可視化
> install.packages("plotrix") # パッケージのインストール
> library(plotrix) # パッケージのロード
> plotCI(1:20, result["xbar", 1:20], ui = result["ui", 1:20],
+        li = result["li", 1:20], pch = 4, axes = FALSE, xlab = "",
+        ylab = expression(mu), lwd = 2) # 信頼区間と標本平均の図示
> axis(1, 1:20, 1:20) # x 軸の追加
> axis(2) # y 軸の追加
> abline(h = mu, col = "red", lty = "dashed", lwd = 2) # 真の平均の図示
```



8.4.2. 平均の区間推定 (分散未知). 分散 σ^2 が既知であることは稀である. σ^2 が未知の場合, 不偏分散 $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ で代用するのが自然な考えである. s^2 の分布については次の結果が知られている ($n \geq 2$ とする):

命題 8.2. X_1, X_2, \dots, X_n は独立同分布な確率変数列で, 平均 μ , 分散 σ^2 の正規分布に従うとする. このとき, \bar{X} と s^2 は独立であり, 確率変数 $(n-1)s^2/\sigma^2$ は自由度 $n-1$ の χ^2 分布に従う.

上の命題と $\sqrt{n}(\bar{X} - \mu)/\sigma$ が標準正規分布に従うことから, 確率変数

$$\frac{\sqrt{n}(\bar{X} - \mu)}{s} \left(= \frac{\sqrt{n}(\bar{X} - \mu)/\sigma}{\sqrt{\frac{(n-1)s^2}{\sigma^2}/(n-1)}} \right)$$

は自由度 $n-1$ の t 分布に従うことがわかる (6.3.5 節参照). 従って, $\alpha \in (0, 1)$ を定めるとき, $t_{1-\alpha/2}(n-1)$ を自由度 $n-1$ の t 分布の $100(1-\alpha/2)\%$ 分位点とすれば, 上と同様の議論によって,

$$P\left(-t_{1-\alpha/2}(n-1) \leq \frac{\sqrt{n}(\bar{X} - \mu)}{s} \leq t_{1-\alpha/2}(n-1)\right) = 1 - \alpha$$

が成り立つことがわかる. カッコ内を μ について解くことで,

$$[\bar{X} - t_{1-\alpha/2}(n-1) \cdot s/\sqrt{n}, \bar{X} + t_{1-\alpha/2}(n-1) \cdot s/\sqrt{n}]$$

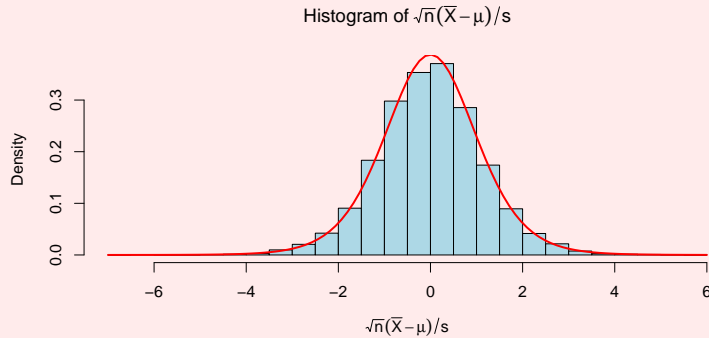
が平均 μ の $100(1-\alpha)\%$ 信頼区間を与えることがわかる.

```
> set.seed(111) # 乱数の初期値の設定
> mu <- 10 # 平均
> sigma <- 5 # 標準偏差
> # n 個の正規標本から標本平均・標準偏差と平均の 100(1-alpha)%信頼区間の
> # 上端・下端を計算する関数の作成
> myest <- function(n, alpha = 0.05){
+
+   x <- rnorm(n, mean = mu, sd = sigma) # 正規乱数の生成
+   xbar <- mean(x) # 標本平均
+   s <- sd(x) # 標準偏差 (不偏分散の平方根)
+   ui <- xbar + qt(1-alpha/2, df = n - 1) * s/sqrt(n) # 上側信頼限界
+   li <- xbar - qt(1-alpha/2, df = n - 1) * s/sqrt(n) # 下側信頼限界
+
+   return(c(xbar = xbar, s = s, ui = ui, li = li))
+ }
```

```

> result <- replicate(10000, myest(10)) # n=10として10000回実験
> ## sqrt(n)*(xbar-mu)/sが自由度n-1のt分布に従うことの確認
> hist(sqrt(10) * (result["xbar", ] - mu)/result["s", ], freq = FALSE,
+      col = "lightblue", breaks = 20,
+      xlab = expression(sqrt(n)*(bar(X)-mu)/s),
+      main = expression(paste("Histogram of ", sqrt(n)*(bar(X)-mu)/s)))
> curve(dt(x, df = 9), add = TRUE, col = "red", lwd = 2)

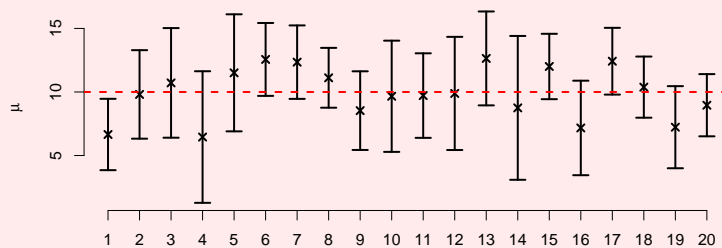
```



```

> ## 真の平均 mu が 95%信頼区間に含まれている割合が約 95%であることの確認
> length(which(mu <= result["ui", ] & mu >= result["li", ]))/10000
[1] 0.95
> ## 最初の 20 個の信頼区間の可視化
> plotCI(1:20, result["xbar",1:20], ui = result["ui",1:20],
+       li = result["li",1:20], pch = 4, axes = FALSE, xlab = "",
+       ylab = expression(mu), lwd = 2) # 信頼区間と標本平均の図示
> axis(1, 1:20, 1:20) # x軸の追加
> axis(2) # y軸の追加
> abline(h = mu, col = "red", lty = "dashed", lwd = 2) # 真の平均の図示

```



(ci-mean-unknown.r)

8.4.3. 分散の区間推定. 分散 σ^2 の区間推定には命題 8.2 を利用する. すなわち, $(n-1)s^2/\sigma^2$ が自由度 $n-1$ の χ^2 分布に従うので, $\chi_{\alpha/2}^2(n-1)$, $\chi_{1-\alpha/2}^2(n-1)$ をそれぞれ自由度 $n-1$ の χ^2 分布の $100\alpha/2\%$, $100(1-\alpha/2)\%$ 分位点とすれば,

$$(8.2) \quad P(\chi_{\alpha/2}^2(n-1) \leq (n-1)s^2/\sigma^2 \leq \chi_{1-\alpha/2}^2(n-1)) = 1 - \alpha$$

が成り立つ。実際、自由度 $n-1$ の χ^2 分布の確率密度関数を f とすれば、

$$\begin{aligned} P(\chi_{\alpha/2}^2(n-1) \leq (n-1)s^2/\sigma^2 \leq \chi_{1-\alpha/2}^2(n-1)) &= \int_{\chi_{\alpha/2}^2(n-1)}^{\chi_{1-\alpha/2}^2(n-1)} f(x)dx \\ &= \int_{-\infty}^{\chi_{1-\alpha/2}^2(n-1)} f(x)dx - \int_{-\infty}^{\chi_{\alpha/2}^2(n-1)} f(x)dx = 1 - \alpha/2 - \alpha/2 = 1 - \alpha \end{aligned}$$

となる。(8.2) の左辺のカッコ内を σ^2 について解くと、

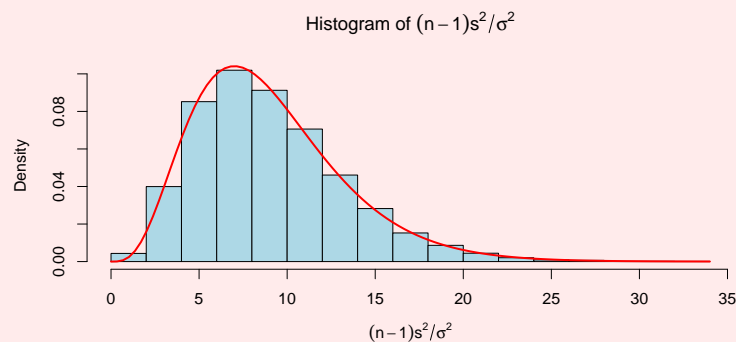
$$P\left((n-1)s^2/\chi_{1-\alpha/2}^2(n-1) \leq \sigma^2 \leq (n-1)s^2/\chi_{\alpha/2}^2(n-1)\right) = 1 - \alpha$$

が得られるので、

$$\left[(n-1)s^2/\chi_{1-\alpha/2}^2(n-1), (n-1)s^2/\chi_{\alpha/2}^2(n-1) \right]$$

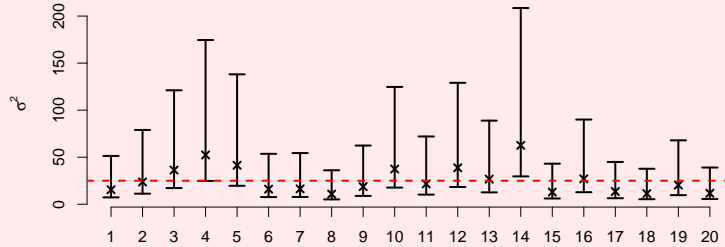
が σ^2 の $100(1-\alpha)\%$ 信頼区間を与える。

```
> set.seed(111) # 乱数の初期値の設定
> mu <- 10 # 平均
> sigma <- 5 # 標準偏差
> # n 個の正規標本から不偏分散と分散の 100(1-alpha)%信頼区間の
> # 上端・下端を計算する関数の作成
> myst <- function(n, alpha = 0.05){
+   x <- rnorm(n, mean = mu, sd = sigma) # 正規乱数の生成
+   s2 <- var(x) # 不偏分散
+   ui <- (n-1) * s2/qchisq(alpha/2, df = n - 1) # 上側信頼限界
+   li <- (n-1) * s2/qchisq(1-alpha/2, df = n - 1) # 下側信頼限界
+   return(c(s2 = s2, ui = ui, li = li))
+ }
> result <- replicate(10000, myst(10)) # n=10として10000回実験
> ## (n-1)*s^2/sigma^2が自由度n-1のカイ二乗分布に従うことの確認
> hist(9 * result["s2", ]/sigma^2, freq = FALSE, col = "lightblue",
+      breaks = 20, xlab = expression((n-1)*s^2/sigma^2),
+      main = expression(paste("Histogram of ", (n-1)*s^2/sigma^2)))
> curve(dchisq(x, df = 9), add = TRUE, col = "red", lwd = 2)
```



```
> ## 真の分散 sigma^2 が 95%信頼区間に含まれている割合が約 95%であることの確認
> length(which(sigma^2 <= result["ui", ] & sigma^2 >= result["li", ]))/10000
[1] 0.9528
> ## はじめの 20 個の信頼区間の可視化
> plotCI(1:20, result["s2", 1:20], ui = result["ui", 1:20],
+       li = result["li", 1:20], pch = 4, axes = FALSE, xlab = "",
+       ylab = expression(sigma^2), lwd = 2) # 信頼区間と不偏分散の図示
> axis(1, 1:20, 1:20) # x 軸の追加
```

```
> axis(2) # y軸の追加
> abline(h = sigma^2, col = "red", lty = "dashed", lwd = 2) # 真の分散の図示
```



(ci-variance.r)

8.5. 漸近正規性による区間推定

前節のように信頼区間を正確に計算できることは稀である。しかし、未知パラメータ θ のある推定量 $\hat{\theta}$ について、推定誤差 $\hat{\theta} - \theta$ の分布がある正規分布で近似できる状況はしばしばある。このような推定量の性質を **漸近正規性** と呼ぶ。漸近正規性をもつ推定量がある場合、推定誤差がある区間に含まれる確率を近似的に求めることができるから、近似的に正しい信頼区間を構成することが可能となる。この節ではそのような方法の具体例を説明する。

8.5.1. 平均の信頼区間. 以前の章で確認したように、確率分布 \mathcal{L} が2次のモーメントを持てば、中心極限定理より、 \mathcal{L} の平均 μ の推定量である標本平均 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ は漸近正規性をもつ。より正確に述べると、 \mathcal{L} の標準偏差を σ とすれば、任意の $a \leq b$ に対して、

$$(8.3) \quad P\left(a \leq \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq b\right) \rightarrow \int_a^b \phi(x) dx \quad (n \rightarrow \infty)$$

が成り立つ (ϕ は標準正規分布の確率密度関数)。従って、 $\alpha \in (0, 1)$ を定めたとき、 $z_{1-\alpha/2}$ を標準正規分布の $100(1-\alpha/2)\%$ 分位点とすれば、

$$P\left(-z_{1-\alpha/2} \leq \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq z_{1-\alpha/2}\right) \rightarrow 1 - \alpha \quad (n \rightarrow \infty)$$

が成り立つ。カッコ内を μ について解くと、

$$P\left(\bar{X} - z_{1-\alpha/2} \cdot \sigma / \sqrt{n} \leq \mu \leq \bar{X} + z_{1-\alpha/2} \cdot \sigma / \sqrt{n}\right) \rightarrow 1 - \alpha \quad (n \rightarrow \infty)$$

が得られるので、 σ が既知であれば、

$$[\bar{X} - z_{1-\alpha/2} \cdot \sigma / \sqrt{n}, \bar{X} + z_{1-\alpha/2} \cdot \sigma / \sqrt{n}]$$

はサンプル数 n が十分大きい場合に近似的に正しい平均 μ の $100(1-\alpha)\%$ 信頼区間を与える。通常は σ は未知であるが、近似 (8.3) は σ をその一致推定量 $\hat{\sigma}$ で置き換えてもそのまま成立することが知られている。従って、上の式で σ を $\hat{\sigma}$ で置き換えたもの

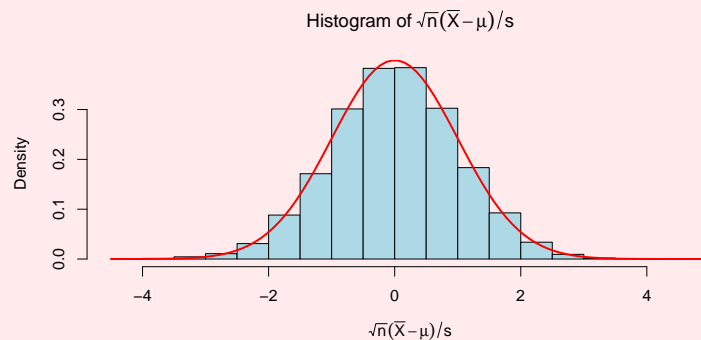
$$[\bar{X} - z_{1-\alpha/2} \cdot \hat{\sigma} / \sqrt{n}, \bar{X} + z_{1-\alpha/2} \cdot \hat{\sigma} / \sqrt{n}]$$

も、サンプル数 n が十分大きい場合に近似的に正しい平均 μ の $100(1-\alpha)\%$ 信頼区間を与える。 $\hat{\sigma}$ としては例えば不偏分散の平方根

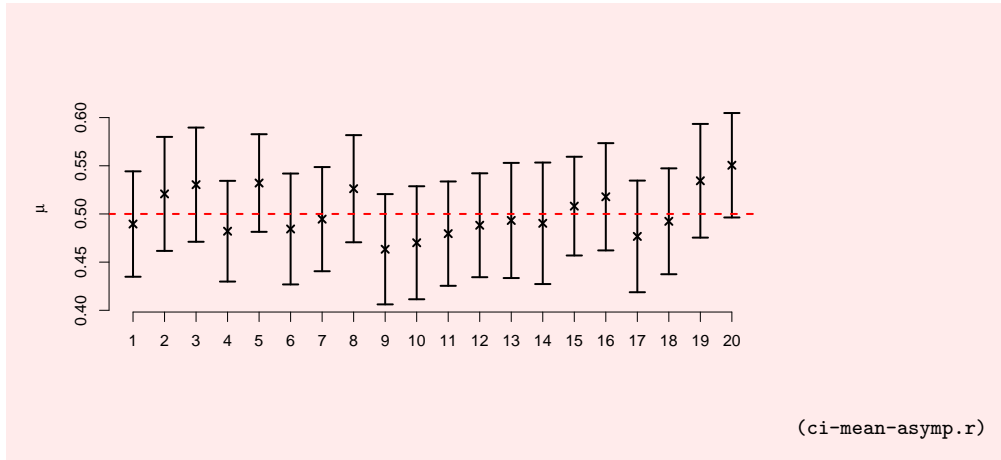
$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

を使うことができる。

```
> ### 一様分布の平均の信頼区間
> set.seed(111) # 乱数の初期値の設定
> # n 個の (0,1) 上の一様乱数から標本平均・標準偏差と平均の 100(1-alpha)%信頼区間の
> # 上端・下端を計算する関数の作成
> myest <- function(n, alpha = 0.05){
+
+   x <- runif(n) # 一様乱数の生成
+   xbar <- mean(x) # 標本平均
+   s <- sd(x) # 標準偏差 (不偏分散の平方根)
+   ui <- xbar + qnorm(1-alpha/2) * s/sqrt(n) # 上側信頼限界
+   li <- xbar - qnorm(1-alpha/2) * s/sqrt(n) # 下側信頼限界
+
+   return(c(xbar = xbar, s = s, ui = ui, li = li))
+ }
> n <- 100 # サンプル数
> result <- replicate(10000, myest(n)) # 10000 回実験
> ## sqrt(n)*(xbar-mu)/s の分布が標準正規分布で近似できることの確認
> mu <- 0.5 # 真の平均
> hist(sqrt(n) * (result["xbar", ] - mu)/result["s", ], freq = FALSE,
+       col = "lightblue", breaks = 20,
+       xlab = expression(sqrt(n)*(bar(X)-mu)/s),
+       main = expression(paste("Histogram of ", sqrt(n)*(bar(X)-mu)/s)))
> curve(dnorm, add = TRUE, col = "red", lwd = 2)
```



```
> ## 真の平均 mu が 95%信頼区間に含まれている割合が約 95%であることの確認
> length(which(mu <= result["ui", ] & mu >= result["li", ]))/10000
[1] 0.9484
> ## はじめの 20 個の信頼区間の可視化
> plotCI(1:20, result["xbar",1:20], ui = result["ui",1:20],
+        li = result["li",1:20], pch = 4, axes = FALSE, xlab = "",
+        ylab = expression(mu), lwd = 2) # 信頼区間と標本平均の図示
> axis(1, 1:20, 1:20) # x 軸の追加
> axis(2) # y 軸の追加
> abline(h = mu, col = "red", lty = "dashed", lwd = 2) # 真の平均の図示
```



8.5.2. 最尤法による区間推定. 広い範囲の確率分布に対して、最尤推定量は漸近正規性をもつことが知られている。より正確に述べるために、確率分布 \mathcal{L} は p 個の未知パラメータ $\boldsymbol{\theta} := (\theta_1, \dots, \theta_p)$ を含む確率分布でモデル化されているとして、その対数尤度関数を $\ell(\boldsymbol{\theta})$ 、最尤推定量を $\hat{\boldsymbol{\theta}}$ とする。また、 $\ell(\boldsymbol{\theta})$ は C^2 級であると仮定し、 p 次正方行列 $I(\boldsymbol{\theta})$ を

$$I(\boldsymbol{\theta}) = \left(\frac{1}{n} E \left[\left(\frac{\partial \ell}{\partial \theta_i}(\boldsymbol{\theta}) \right) \left(\frac{\partial \ell}{\partial \theta_j}(\boldsymbol{\theta}) \right) \right] \right)_{1 \leq i, j \leq p}$$

で定義する (期待値は定義できると仮定する)。さらに、真のパラメータ値を $\boldsymbol{\theta}^* = (\theta_1^*, \dots, \theta_p^*)$ とし、 $I(\boldsymbol{\theta}^*)$ は正則であると仮定する。このとき、緩やかな仮定の下で、各 $j = 1, \dots, p$ について、 v_j を行列 $I(\boldsymbol{\theta}^*)^{-1}$ の第 j 対角成分とすると、推定誤差 $\hat{\theta}_j - \theta_j^*$ の分布は平均 0、分散 v_j/n の正規分布で近似できることが知られている。より正確には、任意の $a \leq b$ に対して、

$$(8.4) \quad P \left(a \leq \frac{\sqrt{n}(\hat{\theta}_j - \theta_j^*)}{\sqrt{v_j}} \leq b \right) \rightarrow \int_a^b \phi(x) dx \quad (n \rightarrow \infty)$$

が成り立つ。従って、上と同様の議論によって、

$$[\hat{\theta}_j - z_{1-\alpha/2} \cdot v_j / \sqrt{n}, \hat{\theta}_j + z_{1-\alpha/2} \cdot v_j / \sqrt{n}]$$

はサンプル数 n が十分大きい場合に近似的に正しいパラメータ θ_j の $100(1-\alpha)\%$ 信頼区間を与える。一般には v_j は未知パラメータ $\boldsymbol{\theta}^*$ に依存するため、 $I(\hat{\boldsymbol{\theta}})^{-1}$ の第 j 対角成分で置き換える (この場合でも近似式 (8.4) は成り立つため、上の議論は引き続き正当化される)。

行列 $I(\boldsymbol{\theta}^*)$ は **Fisher 情報行列** と呼ばれる。特に $p = 1$ の場合は 8.1 節で述べた Fisher 情報量と同じものとなる。多くの場合に次の等式が成り立つ:

$$I(\boldsymbol{\theta}^*) = \left(-\frac{1}{n} E \left[\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j}(\boldsymbol{\theta}^*) \right] \right)_{1 \leq i, j \leq p}.$$

このことから、 $I(\boldsymbol{\theta}^*)$ を解析的に求めるのが困難な場合、

$$\left(-\frac{1}{n} \frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j}(\boldsymbol{\theta}^*) \right)_{1 \leq i, j \leq p}$$

で代用することが多い (上の行列の各成分は独立同分布な確率変数たちの平均で表されるため、適当な仮定の下で大数の強法則より $I(\boldsymbol{\theta}^*)$ の一致推定量となる)。関数 `optim()` で対数尤度関数のマイナスを最小化する場合、オプション `hessian` を `TRUE`

に指定することで、上の行列に n をかけたものを数値計算した値を計算するよう
 できる (詳細は `help(optim)` を参照すること).

例 8.5 (Poisson 分布の Fisher 情報量). 例 8.2 より、パラメーター λ の Poisson 分布の Fisher 情報量は $I(\lambda) = \lambda^{-1}$ で与えられる.

例 8.6 (指数分布の Fisher 情報量). 例 8.3 より、パラメーター λ の指数分布の Fisher 情報量は $I(\lambda) = \lambda^{-2}$ で与えられる.

例 8.7 (ガンマ分布の Fisher 情報行列). 例 8.4 より、パラメーター ν, α のガンマ分布の Fisher 情報行列は

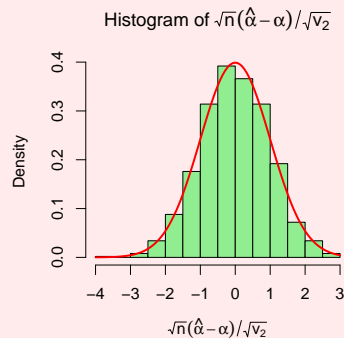
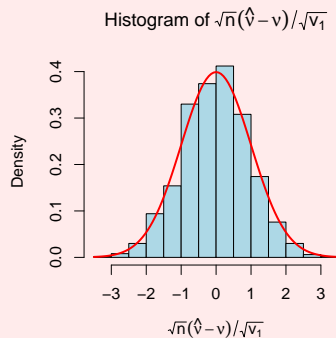
$$I(\nu, \alpha) = \begin{pmatrix} \psi'(\nu) & -\alpha^{-1} \\ -\alpha^{-1} & \nu\alpha^{-2} \end{pmatrix}$$

で与えられる. ここで、ディガンマ関数の導関数 $\psi'(\nu)$ はトリガンマ関数と呼ばれ、R
 では関数 `trigamma()` で計算できる. Fisher 情報行列の逆行列は、

$$I(\nu, \alpha)^{-1} = \frac{1}{\nu\psi'(\nu) - 1} \begin{pmatrix} \nu & \alpha \\ \alpha & \alpha^2\psi'(\nu) \end{pmatrix}$$

となる.

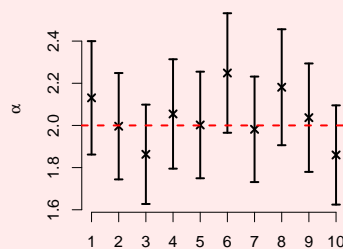
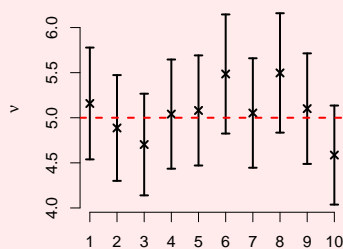
```
> ### 最尤法によるガンマ分布のパラメーターの区間推定
> ## シミュレーションによる検証
> # n 個のガンマ乱数から標本平均・標準偏差と平均の 100(1-alpha)%信頼区間の
> # 上端・下端を計算する関数の作成
> myest <- function(n, shape, rate, alpha = 0.05){
+
+   x <- rgamma(n, shape, rate) # ガンマ乱数の生成
+   thetahat <- mle.gamma(x) # 最尤推定量
+   names(thetahat) <- NULL
+   v1hat <- thetahat[1]/(thetahat[1]*trigamma(thetahat[1])-1) # nu の最尤推定量の漸近分散
+   v2hat <- thetahat[2]^2*trigamma(thetahat[1])/
+     (thetahat[1]*trigamma(thetahat[1])-1) # alpha の最尤推定量の漸近分散
+   ui1 <- thetahat[1] + qnorm(1-alpha/2) * sqrt(v1hat/n) # nu の上側信頼限界
+   li1 <- thetahat[1] - qnorm(1-alpha/2) * sqrt(v1hat/n) # nu の下側信頼限界
+   ui2 <- thetahat[2] + qnorm(1-alpha/2) * sqrt(v2hat/n) # alpha の上側信頼限界
+   li2 <- thetahat[2] - qnorm(1-alpha/2) * sqrt(v2hat/n) # alpha の下側信頼限界
+
+   return(c(nu = thetahat[1], alpha = thetahat[2], v1hat = v1hat, v2hat = v2hat,
+     ui1 = ui1, li1 = li1, ui2 = ui2, li2 = li2))
+ }
> set.seed(123) # 乱数の初期化
> nu <- 5
> alpha <- 2
> n <- 500 # サンプル数
> result <- replicate(1000, myest(n, nu, alpha)) # 1000 回実験
> op <- par(mfcol = c(1, 2)) # 画面を 2 分割
> ## 漸近正規性の確認
> hist(sqrt(n) * (result["nu", ] - nu)/sqrt(result["v1hat", ]), freq = FALSE,
+   col = "lightblue", breaks = 20,
+   xlab = expression(sqrt(n)*(hat(nu)-nu)/sqrt(v[1])),
+   main = expression(paste("Histogram of ", sqrt(n)*(hat(nu)-nu)/sqrt(v[1]))))
> curve(dnorm, add = TRUE, col = "red", lwd = 2)
> hist(sqrt(n) * (result["alpha", ] - alpha)/sqrt(result["v2hat", ]), freq = FALSE,
+   col = "lightgreen", breaks = 20,
+   xlab = expression(sqrt(n)*(hat(alpha)-alpha)/sqrt(v[2])),
+   main = expression(paste("Histogram of ", sqrt(n)*(hat(alpha)-alpha)/sqrt(v[2]))))
> curve(dnorm, add = TRUE, col = "red", lwd = 2)
> par(op)
```



```

> ## 真のパラメータが95%信頼区間に含まれている割合が約95%であることの確認
> length(which(nu <= result["ui1", ] & nu >= result["li1", ]))/1000
[1] 0.956
> length(which(alpha <= result["ui2", ] & alpha >= result["li2", ]))/1000
[1] 0.954
> ## 最初の10個の信頼区間の可視化
> op <- par(mfcol = c(1, 2)) # 画面を2分割
> plotCI(1:10, result["nu", 1:10], ui = result["ui1", 1:10],
+       li = result["li1", 1:10], pch = 4, axes = FALSE, xlab = "",
+       ylab = expression(nu), lwd = 2) # 信頼区間と最尤推定値の図示
> axis(1, 1:10, 1:10) # x軸の追加
> axis(2) # y軸の追加
> abline(h = nu, col = "red", lty = "dashed", lwd = 2) # 真のパラメータの図示
> plotCI(1:10, result["alpha", 1:10], ui = result["ui2", 1:10],
+       li = result["li2", 1:10], pch = 4, axes = FALSE, xlab = "",
+       ylab = expression(alpha), lwd = 2) # 信頼区間と最尤推定値の図示
> axis(1, 1:10, 1:10) # x軸の追加
> axis(2) # y軸の追加
> abline(h = alpha, col = "red", lty = "dashed", lwd = 2) # 真のパラメータの図示
> par(op)

```



```

> ## 実データへの適用
> kikou <- read.csv("kikou2016.csv", fileEncoding = "sjis") # データの読み込み
> x <- kikou$風速 # 風速のデータにガンマ分布をあてはめてみる
> (thetahat <- mle.gamma(x)) # 最尤推定
      nu      alpha
12.233964 4.392418
> v1hat <- thetathat[1]/(thetathat[1]*trigamma(thetathat[1])-1) # nuの最尤推定量の漸近分散
> v2hat <- thetathat[2]^2*trigamma(thetathat[1])/
+ (thetathat[1]*trigamma(thetathat[1])-1) # alphaの最尤推定量の漸近分散

```

```
> ui1 <- thetahat[1] + qnorm(1-0.05/2) * sqrt(v1hat/n) # nu の 95%上側信頼限界
> li1 <- thetahat[1] - qnorm(1-0.05/2) * sqrt(v1hat/n) # nu の 95%下側信頼限界
> c(li1, ui1) # nu の 95%信頼区間
      nu      nu
10.73767 13.73025
> ui2 <- thetahat[2] + qnorm(1-0.05/2) * sqrt(v2hat/n) # alpha の 95%上側信頼限界
> li2 <- thetahat[2] - qnorm(1-0.05/2) * sqrt(v2hat/n) # alpha の 95%下側信頼限界
> c(li2, ui2) # alpha の 95%信頼区間
      alpha    alpha
3.844037 4.940799
```

(ci-mle.r)

演習 8.2. 最尤推定量の漸近正規性について調べてみよう.

- (1) 上の実行例において, 最尤推定量の漸近分散を関数 `optim()` のオプション `hessian` を `TRUE` と指定して計算する方法について調べてみよ.
- (2) Poisson 分布や指数分布の場合の最尤推定量の漸近正規性について, シミュレーションで確認せよ.

8.6. 参考文献

1. 東京大学教養学部統計学教室編「統計学入門」, 東京大学出版会 (1991 年).
2. 吉田朋広著「数理統計学」, 朝倉書店 (2006 年).