rmarkdown パッケージで楽々ドキュメント生成

@kohske

June 4, 2014

はじめに

Rマークダウンでドキュメントとコード書いてrmarkdown::render() します。

あやめとは、

▶ さかな植物の名前です。

- さかな植物の名前です。
- ▶ おそらく、世界中でも最も多く解析にさらされた植物でしょう。

- さかな植物の名前です。
- ▶ おそらく、世界中でも最も多く解析にさらされた植物で しょう。
- ▶ 学名は Iris sanguinea といいます。

- さかな植物の名前です。
- ▶ おそらく、世界中でも最も多く解析にさらされた植物でしょう。
- ▶ 学名は Iris sanguinea といいます。
- ▶ イリスではなくて、アイリスです。

- さかな植物の名前です。
- ▶ おそらく、世界中でも最も多く解析にさらされた植物で しょう。
- ▶ 学名は Iris sanguinea といいます。
- ▶ イリスではなくて、アイリスです。
- 大きい声では言えませんが今でも「イリス」と呼んでます。

データの雰囲気

ここでは先頭の6行を見てみましょう。

knitr::kable(head(iris), format = "pandoc", caption="あやめのデータ

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa

Table 1: あやめのデータ (1-6 行)

データの解析

変数間の相関を調べてみましょう。

knitr::kable(cor(iris[, -5]), format = "pandoc", caption="あやめの相

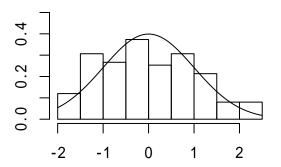
	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length	1.0000	-0.1176	0.8718	0.8179
Sepal.Width	-0.1176	1.0000	-0.4284	-0.3661
Petal.Length	0.8718	-0.4284	1.0000	0.9629
Petal.Width	0.8179	-0.3661	0.9629	1.0000

Table 2: あやめの相関

データの可視化

ヒストグラムを作って、正規分布 $(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right))$ と比べてみます。

par(mar=c(2.5, 2.5, 1.5, 1)) hist(scale(iris[, 1]), probability = TRUE, ylim=c(0, 0.5), main = NU curve(dnorm(x), add=TRUE)



最後に

Enjoy!!