

R と RStudio の使い方

芳賀敏郎 (2011) 医薬品開発のための統計解析 第1部 基礎

1 統計の基礎

1.3 分散の加法性、中心極限定理、正規分布

テキストと利用上の注意

●テキスト

芳賀敏郎（2011）医薬品開発のための統計解析

第1部 基礎 改訂版、サイエンティスト社、p.275

（サイトへアップすることに対して、サイエンティスト社の了解を得ています）

●Rによる解析事例を紹介

R スクリプトの出力結果を紹介します（tidyverse 系には次期バージョンで対応します）

R スクリプト（文字コードUTF-8に設定）を、このサイトからダウンロードできます

R スクリプトを [Compile Report] することにより、Word または HTML で見ることができます

R と RStudio の設定と基本的な使い方は「[R と RStudio の使い方入門](#)」を参照してください

R の出力結果の見方は、テキストとそれを解説した [PDF ファイル](#) を参照してください

グラフ表示は、解析手段として、必要最小限の表現に止めています

●自己責任で利用

上記のことを理解した上で、自己責任により利用してください

第1部 基礎

- 1. 統計の基礎
 - 1.1 宝くじの期待値と分散、1.2 サイコロの目の数の期待値と分散
 - 1.3 分散の加法性・中心極限定理・正規分布**、1.4 統計的推測、1.5 モデル
- 2. 1組のデータの解析
 - 2.1 データの特徴の記述、2.2 データのグラフ表示と外れ値
 - 2.3 対数変換と対数正規分布、2.4 平均に関する推測（母標準偏差 σ 既知）
 - 2.5 分散に関する推測、2.6 平均に関する推測（母標準偏差 σ 未知）
- 3. 2組のデータの解析
 - 3.1 データのグラフ化、3.2 平均値の差の t 検定、3.3 分散の違いの検定
 - 3.4 分散が異なる場合の平均値の差の比較
 - 3.5 対応のある場合の平均値の差の t 検定、3.6 検出力と n の決め方
 - 3.7 ノンパラメトリック検定
- 4. 相関・回帰
 - 4.1 散布図、4.2 相関係数、4.3 回帰モデルとモデルの推定
 - 4.4 誤差を考慮した推定、4.5 回帰分析適用上の諸問題



2つのサイコロの目の合計の分布

● 表示1.3.1 T₂ の分布

(1) 2つのサイコロの目の組み合わせを列挙

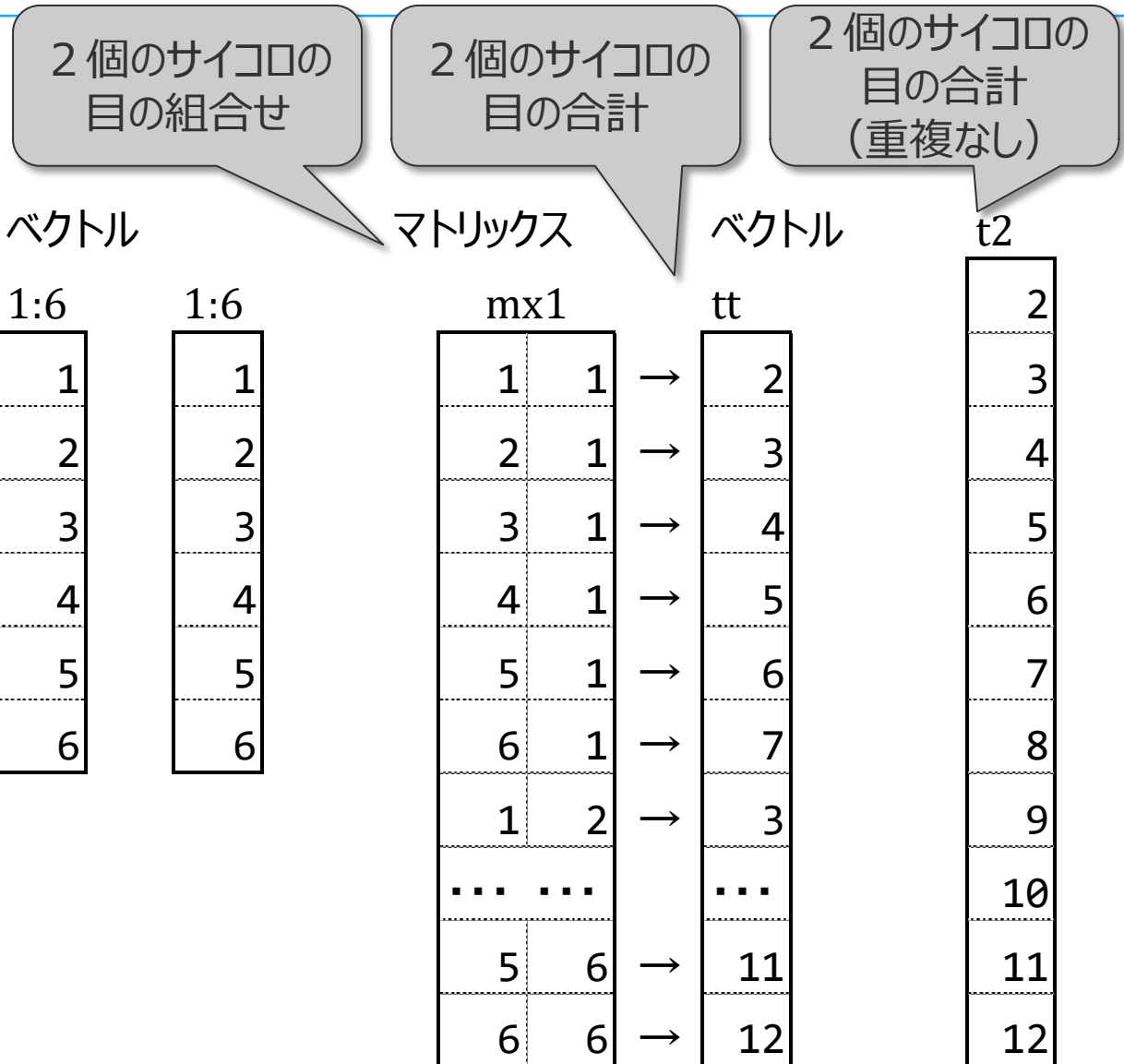
```
mx1 <- as.matrix(
  expand.grid(1:6, 1:6))
```

(2) 2個の目の合計

```
tt <- rowSums(mx1)
```

(3) 2個の目の合計 (重複なし)

```
t2 <- unique(tt)
```





2つのサイコロの目の合計の分布

●表示1.3.1 T₂ の分布

(4) tt (2個のサイコロの目の合計) の
組合せの数をカウントしてnnに付値

```
for (i in 1:length(t2)) {  
  nn[i] <- sum(tt == t2[i])  
}
```

length(t2)
は11

イコール2個

sum(ベクトルと条件)
ベクトルの中から条件を
満たす要素数を取得

2個のサイコロの
目の合計
36通り

2個のサイコロの
目の合計
(重複なし)

ttの中で
t2に一致する
要素の個数

tt	1	2	3	4	5	6
	2	3	4	5	6	7
	7	8	9	10	11	12
	3	4	5	6	7	8
	13	14	15	16	17	18
	4	5	6	7	8	9
	19	20	21	22	23	24
	5	6	7	8	9	10
	25	26	27	28	29	30
	6	7	8	9	10	11
	31	32	33	34	35	36
	7	8	9	10	11	12

t2	1	2	3	4	5	6	7	8	9	10	11	12
		2										
i	2	3										
	3	4										
	4	5										
	5	6										
	6	7										
	7	8										
	8	9										
	9	10										
	10	11										
	11	12										

nn	1	2	3	4	5	6	7	8	9	10	11	12
	1											
i	2	2										
	3	3										
	4	4										
	5	5										
	6	6										
	7	5										
	8	4										
	9	3										
	10	2										
	11	1										



2つのサイコロの目の合計の分布

●表示1.3.1 T₂ の分布

スクリプトファイル

Green1-1-3a.R

利用した関数

cbind、rowSums、

colSums、colnames

expand.grid、sort、

cat

方法

expand.grid 関数で

2つのサイコロの組合せ

を作って集計

##	T2	comb	piT	piT*T2	var1	var2	var3	
##	2	1	0.02777778	0.05555556	-5	25	0.6944444	
##	3	2	0.05555556	0.16666667	-4	16	0.8888889	
##	4	3	0.08333333	0.33333333	-3	9	0.7500000	
##	5	4	0.11111111	0.55555556	-2	4	0.4444444	
##	6	5	0.13888889	0.83333333	-1	1	0.1388889	
##	7	6	0.16666667	1.16666667	0	0	0.0000000	
##	8	5	0.13888889	1.11111111	1	1	0.1388889	
##	9	4	0.11111111	1.00000000	2	4	0.4444444	
##	10	3	0.08333333	0.83333333	3	9	0.7500000	
##	11	2	0.05555556	0.61111111	4	16	0.8888889	
##	12	1	0.02777778	0.33333333	5	25	0.6944444	
##	Totals	NA	NA	1.00000000	7.00000000	NA	NA	5.8333333
##	V[T2] = 5.833333							
##	D[T2] = 2.415229							

中心極限定理

p.25

●表示1.3.3 中心極限定理 (サイコロの目の合計の分布)

スクリプトファイル

Green1-1-3b.R

利用した関数

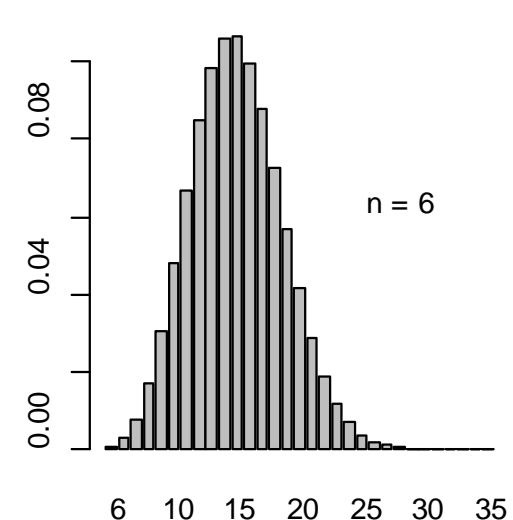
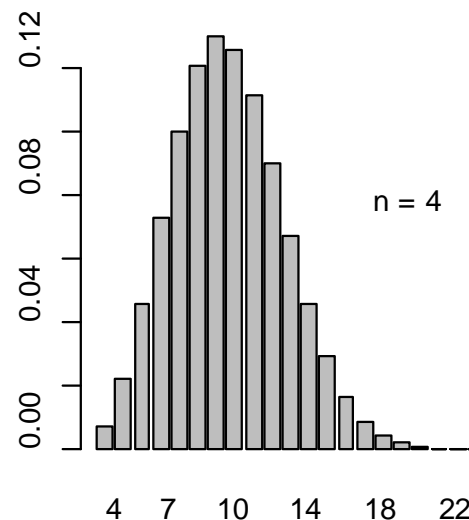
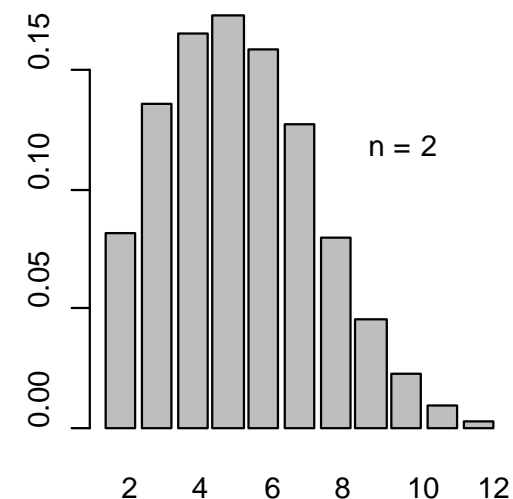
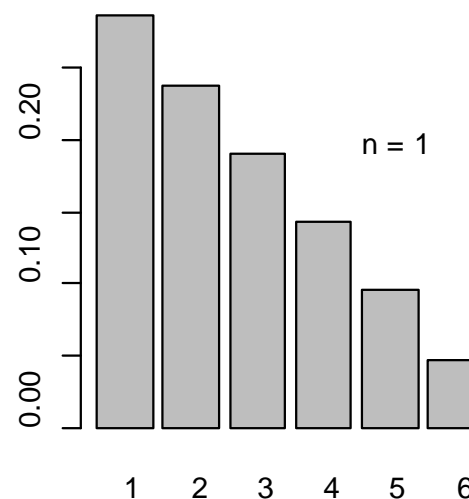
c、matrix、rev、sum、round

par、barplot、text

for

方法

マトリックスからサイコロの目の出る分布を
計算して、棒グラフを作成



正規分布

● 表示1.3.4 正規分布 $N(50, 10^2)$

スクリプトファイル

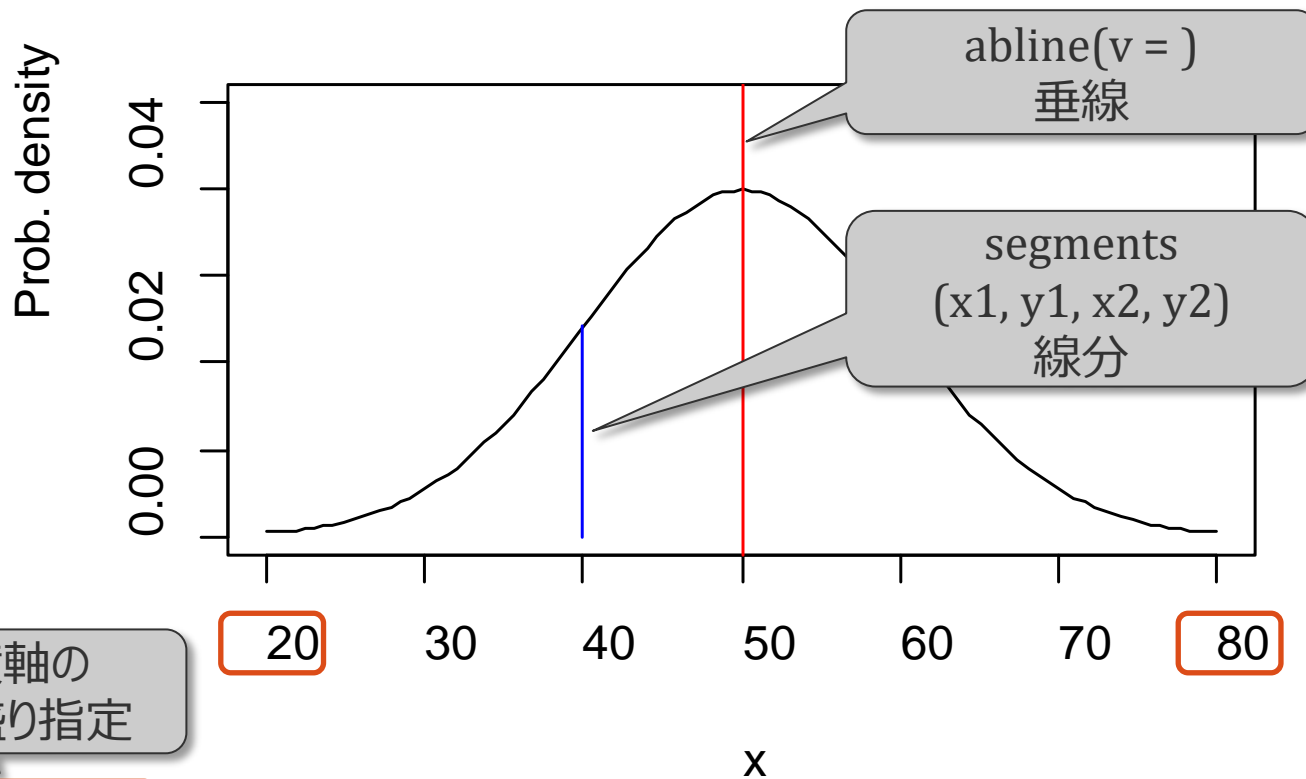
Green1-1-3a.R

利用した関数

curve、dnorm、abline、segments

方法

curve 関数により正規分布を描画



正規分布の
確率密度

x の関数
として記述

横軸の
目盛り指定

```
curve(dnorm(x, mean = 50, sd = 10), 20, 80,  
      ylim = c(0, 0.05), ylab = "Prob. density")
```

```
abline(v = 50, col = 'red', lwd = 1, lty = 1)
```

```
segments(40, 0, 40, dnorm(40, 50, 10), col = "blue", lwd = 1)
```

垂線を付加

線分を付加

x1, y1, x2, y2



- 表示1.3.5 正規分布に関するExcel 関数

- 表示1.3.6 標準正規分布に関するExcel関数

- スクリプトファイル

- Green1-1-3a.R

- 利用した関数

- dnorm

- pnorm

- qnorm

- ifelse

- 下側確率

- `p <- pnorm(x, mean = 50, sd = 10, lower.tail = TRUE)`

- `p <- dnorm(x, mean = 0, sd = 1, lower.tail = TRUE)` # 標準正規分布

- `p <- dnrom(x)` # 標準正規分布

- 上側確率

- `p <- pnorm(x, mean = 50, sd = 10, lower.tail = FALSE)`

- 下側確率が p である x の値

- `q <- qnorm(p, mean = 50, sd = 10, lower.tail = TRUE)`

- 確率密度

- `d <- dnorm(x, mean = 50, sd = 10)`

- `ifelse(条件式, 条件式が成立する場合の結果, 成立しない場合の結果)`



- 表示1.3.5 正規分布に関するExcel 関数

表示1.3.6 標準正規分布に関するExcel関数

スクリプトファイル

Green1-1-3a.R

利用した関数

dnorm、pnorm、qnorm、ifelse、c、cbind

```
##          u Prob density Lower prob Upper prob Tow-sided prob
## [1,] -1  0.241970725  0.1586553 0.841344746  0.317310508
## [2,]  0  0.398942280  0.5000000 0.500000000  1.000000000
## [3,]  1  0.241970725  0.8413447 0.158655254  0.317310508
## [4,]  2  0.053990967  0.9772499 0.022750132  0.045500264
## [5,]  3  0.004431848  0.9986501 0.001349898  0.002699796

##          Lower prob=0.05 Upper Prob=0.05 Two-sided prob=0.05
## [1,]          -1.644854          1.644854          1.959964
```



- 作成 片瀬雅彦
- 作成時期 2021年4月15日