確率・統計
第3回 確率と確率分布

兵庫県立大学　社会情報科学部
川嶋宏彰
kawashima@sis.u-hyogo.ac.jp

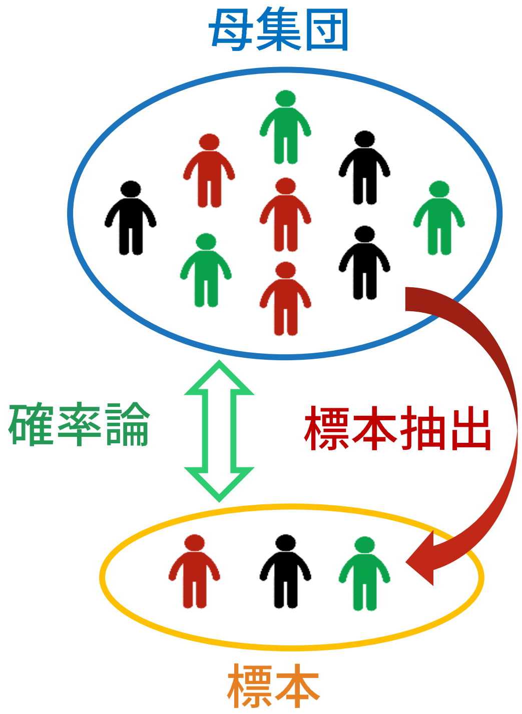
本日の講義内容

- テキスト
  - 「統計学入門」4〜5章
- 確率，確率変数，確率分布
  - [Rcmdr] 確率分布に従う乱数を生成してみる
- 期待値の定義，線形性，平均と分散の加法性
  - 次回よく使う
- 上側（下側，両側）確率，パーセント点
  - [Rcmdr] 正規分布の上側確率やパーセント点を表示してみる
- 宿題（兼レポート予告）

今日は推測統計の準備です

記述統計と推測統計

- 統計には「記述統計」と「推測統計」がある
  - 記述統計 (descriptive statistics)
    - データを要約・視覚化することで理解
    - 平均, 中央値, 分散, 標準偏差, 相関係数, ヒストグラム，散布図
  - 推測統計 (inferential statistics) （推計統計）
    - サンプリング調査を前提（右図）
    - 部分から全体を知る
    - 仮説が正しいかを判断する（仮説検定）
    - 母集団の統計量を見積もる（推定）
    - 過去から未来を予測する



確率論と統計学

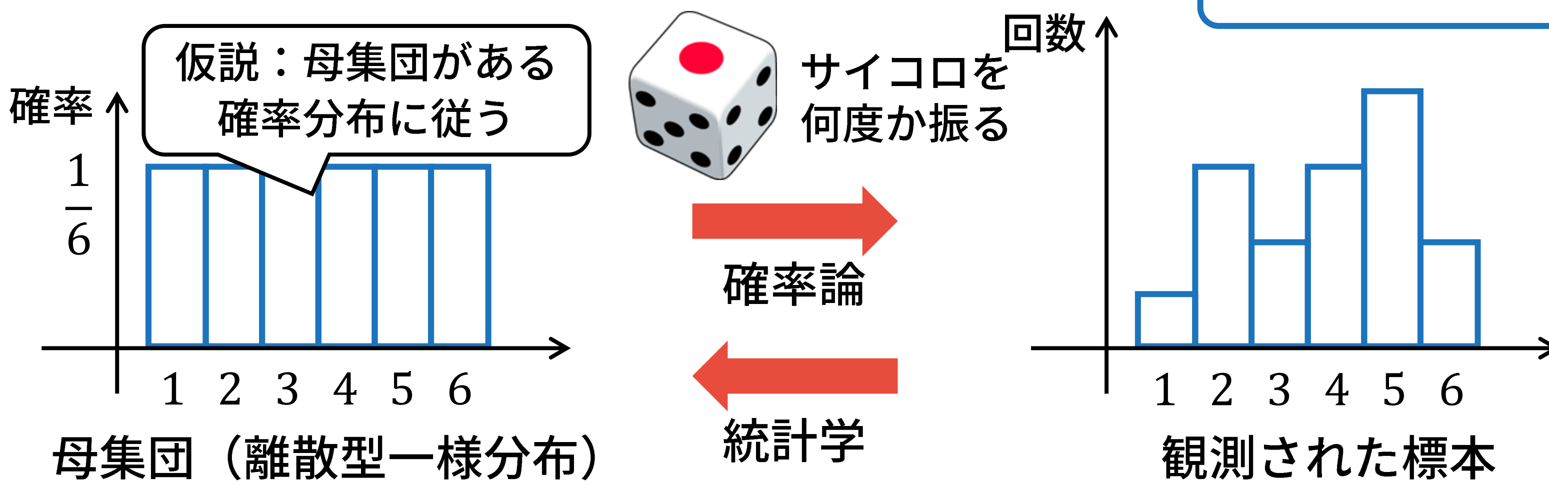
- 統計学（帰納的）
  - 限られた標本からその背後にある母集団を推測したい
- 確率論（演繹的）
  - 仮想的母集団 (確率分布) から標本の得られる確率を計算できる

どうやって？

母集団について
何か仮定する

ここで確率論を使う

確率変数が離散型の場合

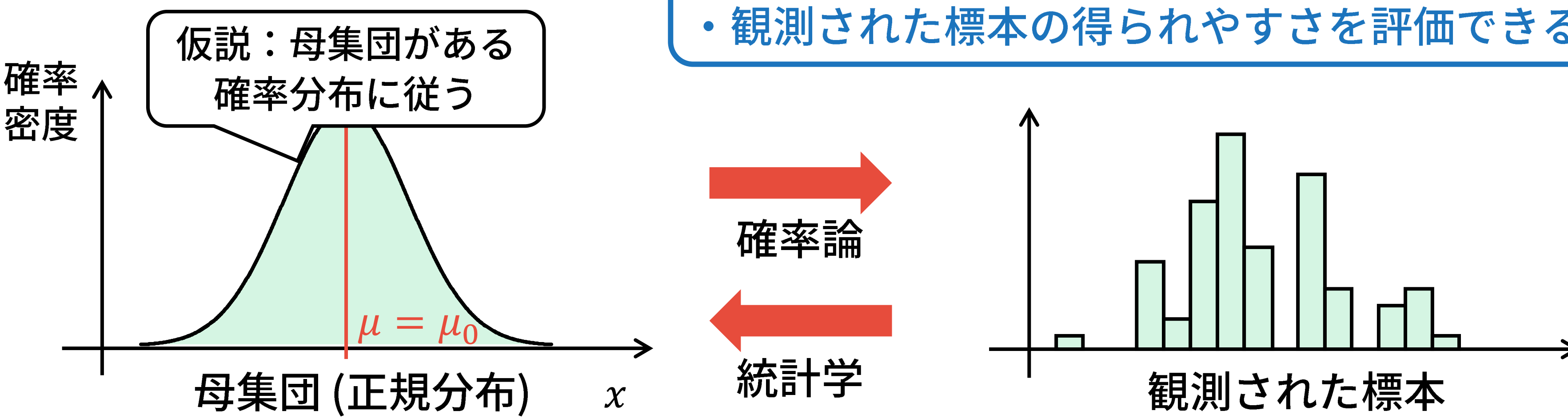


確率論と統計学

- 統計学（帰納的）
  - 限られた標本からその背後にある母集団を推測したい
- 確率論（演繹的）
  - 仮想的母集団 (確率分布) から標本の得られる確率を計算できる

確率変数が連続型の場合

- 仮想的に標本を作ることができる
- 観測された標本の得られやすさを評価できる



事象・排反事象

- 事象 (event)
  - 試行によって起こりうる結果をいくつか集めた集合
  - 例：（試行）サイコロを投げる
    - 1が出る，2が出る，1または2が出る，3以下が出る，4以上が出る，奇数が出る，1から6いずれかが出る

これらはすべて事象

- 標本空間（全事象）Ω
  - 試行によって起こりうる結果のすべての集合
  - 事象は標本空間の部分集合

Q. サイコロ投げでは？

$$\Omega = \{1, \dots, 6\}$$

「1が出る」「2が出る」…を，ここでは省略して，結果の目（数）だけ書くことにする

- 排反事象
  - 共通部分を持たない（同時に起こらない）事象
    - $\{1\}$，$\{2\}, \dots, \{6\}$
    - 奇数が出る $\{1, 3, 5\}$，偶数が出る $\{2, 4, 6\}$

確率

確率：以下の3つ（公理）を満たす数

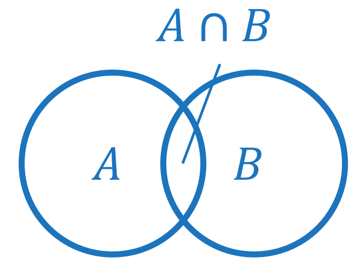
高校では「確率の性質」と呼んでいたもの

1. すべての事象Aに対して$0 \le P(A) \le 1$
2. $P(\Omega) = 1$
3. 排反事象$A_1, A_2, \ldots$に対して$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$

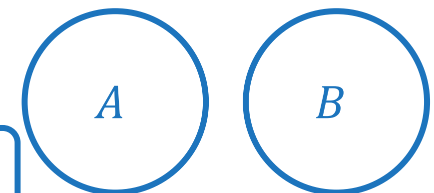
- 例：1つのサイコロを振って出る目
  - $P(\{1\}) = P(\{2\}) = \cdots = 1/6$, $P(\{1,2,3\}) = 1/2$
  - 1から6のいずれかの目が出る確率： $P(\{1,2,3,4,5,6\}) = 1$
  - 3以下の目が出る確率： $P(\{1,2,3\}) = P(\{1\}) + P(\{2\}) + P(\{3\})$

確率の加法定理と排反事象

- 2つの事象 A, B を考える



- 加法定理
  - 一般: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - 事象A, Bが互いに排反（排反事象, $A \cap B = \emptyset$）であるとき
    - $P(A \cup B) = P(A) + P(B)$　（確率の公理）
    - $P(A \cap B) = P(\emptyset) = 0$



A, Bは同時に起こらない

A: 3以下　　$A = \{1, 2, 3\}$
B: 偶数　　$B = \{2, 4, 6\}$

排反でない

$$P(A \cup B) = \frac{3}{6} + \frac{3}{6} - \frac{1}{6}$$

$A = \{1, 2, 3\}$
$B = \{5, 6\}$

排反である

$$P(A \cup B) = \frac{3}{6} + \frac{2}{6}$$

条件付き確率

- 乗法定理

$$P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$$

- $P(B) \neq 0$ なら条件付き確率 $P(A|B)$ は

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- $P(A) \neq 0$ なら条件付き確率 $P(B|A)$ は

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

確率の乗法定理と独立事象（重要, p.82-83）

- Aが起きたと知ったときBについて何か分かるか？
  - 一般: $P(A\cap B) = P(A\mid B)P(B) = P(B\mid A)P(A)$
  - 事象A, Bが互いに独立（独立事象）であるとは
    - $P(A\cap B) = P(A)P(B)$（独立性の定義）

      ($P(A\mid B) = P(A),\ P(B\mid A) = P(B)$ ⋯(☆))

表（カルノー図）に場合の数を入れよう

	B	$\bar{B}$
A	$A\cap B$	$A\cap\bar{B}$
$\bar{A}$	$\bar{A}\cap B$	$\bar{A}\cap\bar{B}$

$A=\{1,2,3\}, B=\{2,4,6\}$ 独立?

	B: 偶数	$\bar{B}$: 奇数	
A $\{1,2,3\}$	$\{2\}$ 1	$\{1,3\}$ 2	3
$\bar{A}$ $\{4,5,6\}$	$\{4,6\}$ 2	$\{5\}$ 1	3
	3	3	6

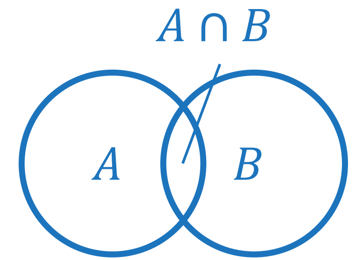
$A=\{1,2\}, B=\{2,4,6\}$ 独立?

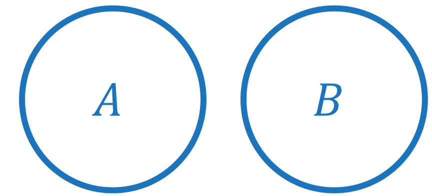
	B: 偶数	$\bar{B}$: 奇数	
A $\{1,2\}$	$\{2\}$ 1	$\{1\}$ 1	2
$\bar{A}$ $\{3,4,5,6\}$	$\{4,6\}$ 2	$\{3,5\}$ 2	4
	3	3	6

確率の加法定理と乗法定理（まとめ）

- 加法定理
  - 一般: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - 事象A, Bが互いに排反 $(A \cap B = \emptyset)$ であるとき
    - $P(A \cup B) = P(A) + P(B)$ （確率の公理）
    - $P(A \cap B) = P(\emptyset) = 0$

同時に起こらない





- 乗法定理（通常排反でない場合を考える）
  - 一般: $P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$
  - 事象A, Bが互いに独立であるとは
    - $P(A \cap B) = P(A)P(B)$ （独立性の定義）

    $(P(A \mid B) = P(A),\ P(B \mid A) = P(B))$

	A	$\bar{A}$
B	$A \cap B$	$\bar{A} \cap B$
$\bar{B}$	$A \cap \bar{B}$	$\bar{A} \cap \bar{B}$

一方を知っても他方の情報が得られない

事象の独立性に関する練習問題

1. 2つ前のスライドの例で条件付き確率を求め，(☆)が成立するか否かを確認せよ．
2. 52枚のトランプから1枚引く試行で「1(エース)が出る」事象と「スペードが出る」事象は独立か？
   - 式を用いて理由も説明すること
   - トランプにジョーカーは入っていないとする
3. サイコロの1から m を赤色に，$m+1$ から6を青色にする．「赤が出る」事象と「偶数が出る」事象が独立になるようなmを求めよ．

確率変数

- 確率変数 (random variable)
  - とりうる各値に対して確率が与えられている変数
  - 確率的に（ランダムに）値が定まる
  （大文字がよく用いられる．Xなど）
- 実現値
  - 確率変数が実際に取る値（観測値）
  （小文字がよく用いられる．xなど）
- 確率変数が○○となる確率
  - 確率変数Xがx以上の値を取る確率: $P(X \geq x)$ $(= P(\{\omega \mid X(\omega) \geq x\})$
  - 確率変数Xが値xを取る確率: $P(X = x)$ $(= P(\{\omega \mid X(\omega) = x\})$

（参考: この講義ではスルーしてよい）
厳密には、確率変数Xは「表が出る」といった数値ではない結果に値（通常は実数値）を割り当てる写像 $X: \Omega \to R$ である．
これで $P(\{表が出る\})$ を $P(X = 1)$
と表せるようになる．つまり$P(X = x)$は
$P(\{\omega \mid X(\omega) = x\})$ の省略形．
（ここで，事象 $\{\omega \mid X(\omega) = x\}$ は、値xが割り振られた結果ω の集合のこと）

確率分布

- 確率変数には離散型と連続型がある
- 離散型の確率変数 X
  - 可算集合$\{x_1, x_2, \ldots\}$の中の一つの値を取る
    - $X \in \{x_1, x_2, \ldots\}$（意味: X は $\{x_1, x_2, \ldots\}$ のいずれかの要素を取る）
    - $\{x_1, x_2, \ldots, x_K\}$（有限集合）の場合も多い
- 連続型の確率変数 X
  - 連続値を取る

    （例）実数を取る確率変数 $X \in \boldsymbol{R}$

離散分布

- **離散型の確率分布（離散分布）とは
離散型の確率変数がとりうる各値に，確率を割り当てたもの**
  - $P(X = x_k) = f(x_k)$ ただし $\sum_{k=1}^{K} f(x_k) = 1$ （有限でないときはK が ∞）
    - 添字無しで「$P(X = x) = f(x)$ ただし $\sum_x f(x) = 1$ 」と書くことも

- 例: サイコロの出る目を確率変数Xで表す
  - $P(X = 1) = P(X = 2) = \cdots = \frac{1}{6}$

x	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

- 例: コインを投げて表が出た時に1，裏が出た時に0をとる確率変数X
  - $P(X = 1) = p,\ P(X = 0) = 1 - p$
（ただし $0 \leq p \leq 1$）

x	1	0
$f(x)$	p	$1 - p$

是非知っておきたい離散分布

- ベルヌーイ (Bernoulli) 分布
  - ベルヌーイ試行1回を行うときの分布

x	1	0
$f(x)$	p	$1-p$

- ベルヌーイ試行
  - コイン投げのように、毎回2種類いずれかの結果をとり、かつそれらの起こる確率がどの回も同じである独立試行
  - 一方の結果が出る確率: $P(X=1)=p$
  - 他方の結果が出る確率: $P(X=0)=1-p$

二項分布

- 二項分布
  - ベルヌーイ試行をn回繰り返したとき，一方がk回（他方が$n-k$回）生起する確率の分布

（例） 1回コインを投げたとき，表が出る確率をpとする

- n回コインを投げたとき，表が出る回数Xの従う分布

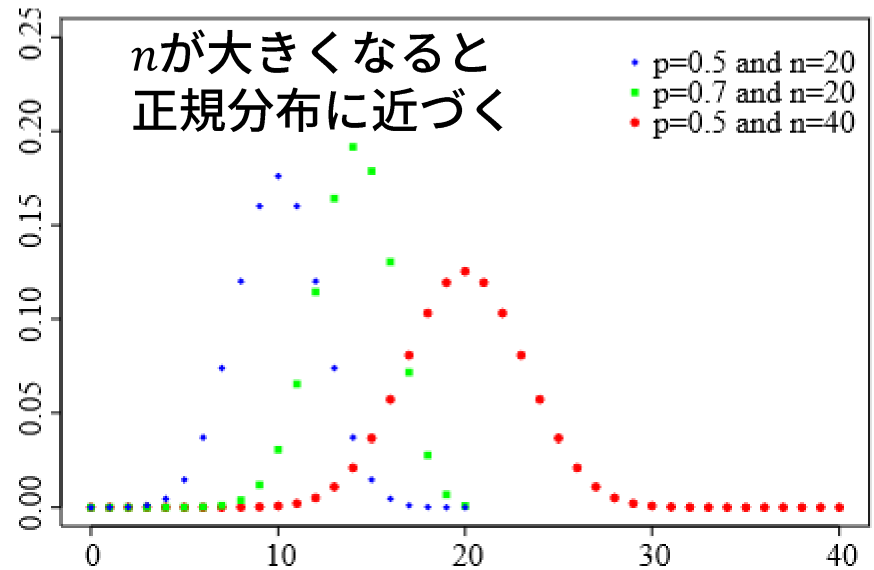
$$P(X=k)={}_nC_k\, p^k(1-p)^{n-k}$$

$n=2$ のとき

k	0	1	2
$f(x_k)$	$(1-p)^2$	$2p(1-p)$	p^2

$n=3$ のとき

k	0	1	2	3
$f(x_k)$	$(1-p)^3$	$3p(1-p)^2$	$3p^2(1-p)$	p^3



連続型の確率分布と確率密度関数

- 連続な値をとる確率変数の分布は密度関数で表す
  - 確率密度 ≠ 確率 であることに注意！

160cm以上，170cm未満の確率は…

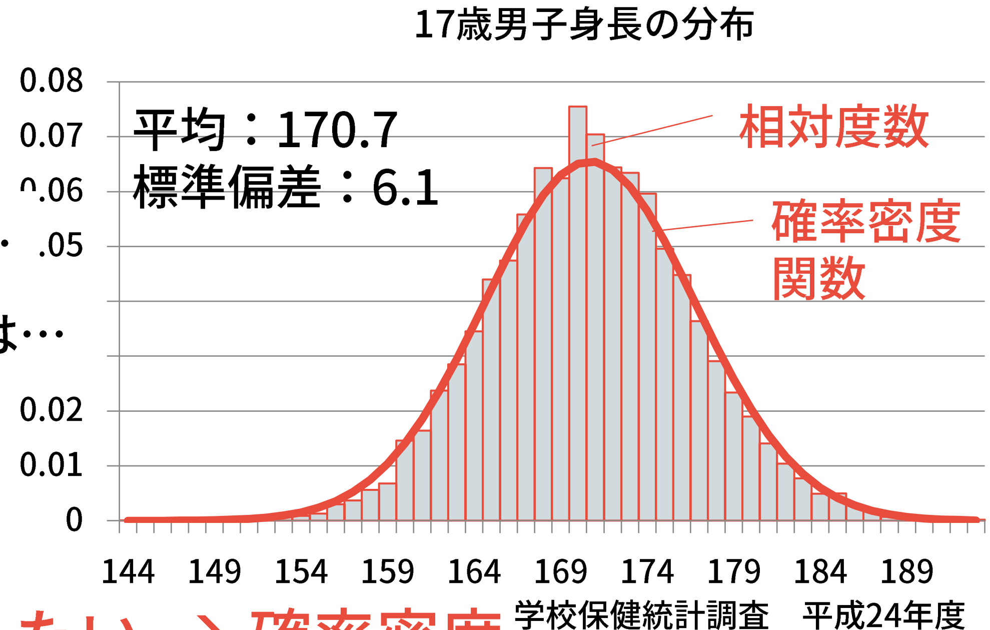
160cm以上，161cm未満の確率は…

160.0cm以上，160.1cm未満の確率は…

160.00cm以上，160.01cm未満の確率は…

身長は元々実数…

どの幅で分布を考えればよい？



十分細かな幅で確率を考えたい → 確率密度

確率と確率密度

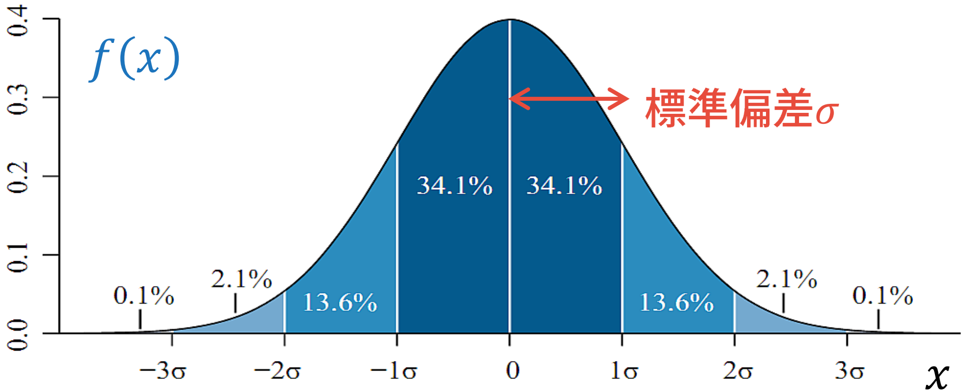
• 連続型の確率分布において：確率密度の積分＝確率

- 連続変数が範囲$[a, b]$の値をとる確率：$P(a \le X \le b) = \int_a^b f(x)dx$

（連続変数がある特定の値aをとる確率：$P(X = a) = 0$）

例：平均0の正規分布

この分布の密度関数は $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$



$P(-\sigma \le X \le \sigma) = 0.6826$

$P(-2\sigma \le X \le 2\sigma) = 0.9544$

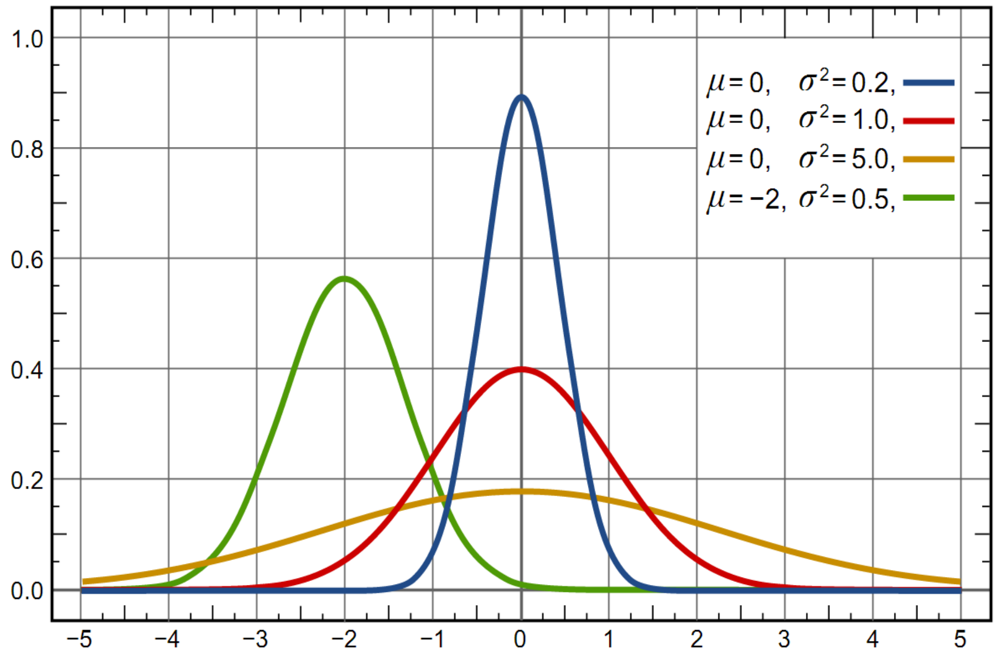
$P(-3\sigma \le X \le 3\sigma) = 0.9974$

確率変数Xの分布がこのような密度関数で記述されるとき，
Xは正規分布$N(0, \sigma^2)$に「従う」といい$X{\sim}N(0, \sigma^2)$のように表記する

正規分布 (Normal distribution)

- 量的な確率変数に関する最も基本的な確率分布の一つ
  - 正規分布$N(\mu, \sigma^2)$に従うデータ:
    平均値 μ を中心に分布 (標準偏差σ はバラつき具合)

確率密度 $f(x)$



x

$N(\mu, \sigma^2)$の確率密度関数

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

密度関数は一般に以下を満たす

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

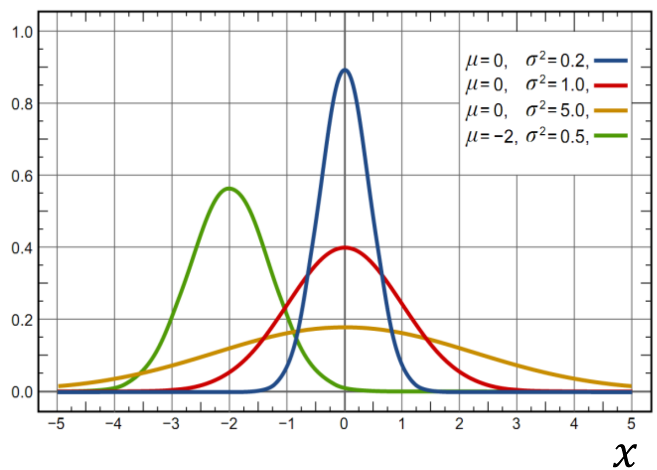
(釣鐘型の面積が1)

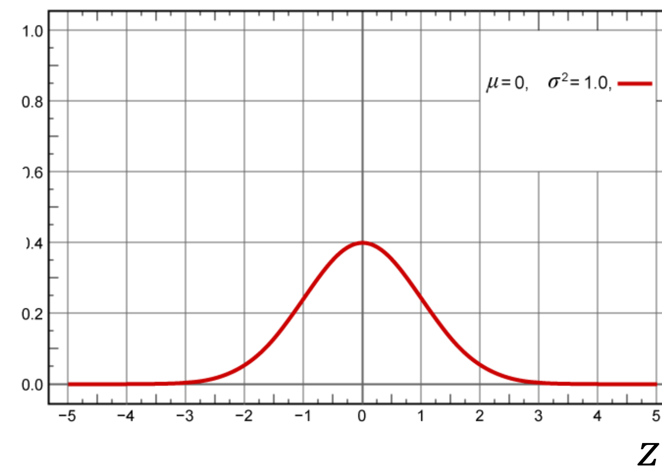
$f(x)$の値自体は1を超えることもある

標準正規分布

- $N(\mu, \sigma^2)$に従う確率変数X → 標準正規分布に従う確率変数Z
- 標準化（変数変換）： $Z = \dfrac{X-\mu}{\sigma}$ （とても重要！）
  - 確率変数Zは平均0・標準偏差1の標準正規分布 $N(0,1)$に従う

確率密度関数: $f_X(x) = \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right)$ ➡ $f_Z(z) = \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{z^2}{2}\right)$

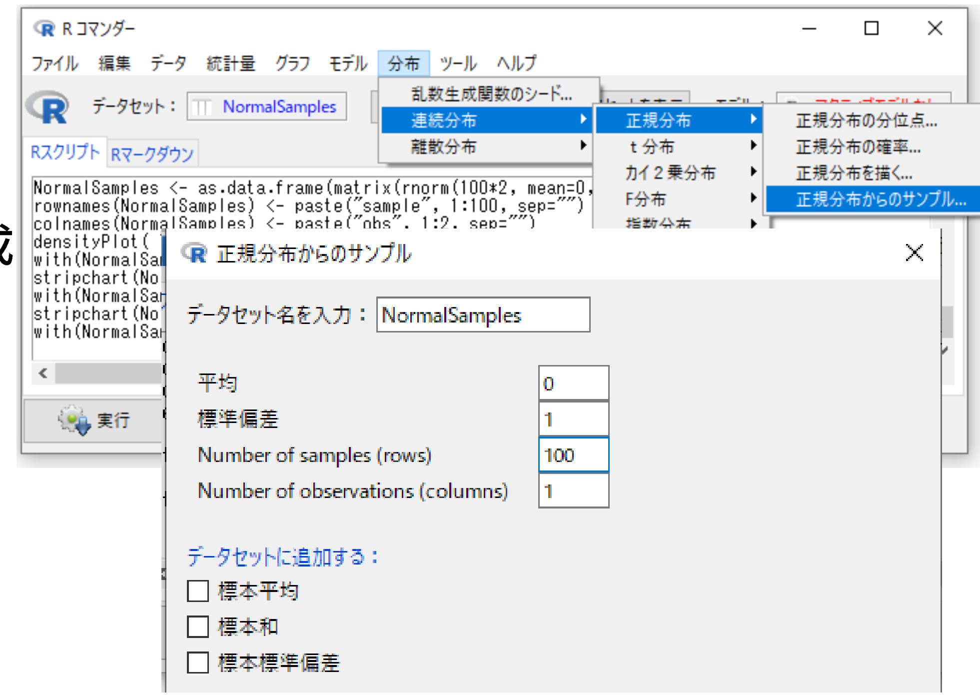




[Rcmdr] 正規乱数を生成

時間が余った人は「離散分布」も試す

- 標準正規分布に従う乱数を生成
  - 分布 > 連続分布 > 正規分布 > 正規分布からのサンプル
  - 平均: 0，標準偏差: 1
  - Number of samples: 100（行数）
  - Number of observations: 1（列数）
    - ヒストグラム等でプロットするために，1列に100個の値を生成

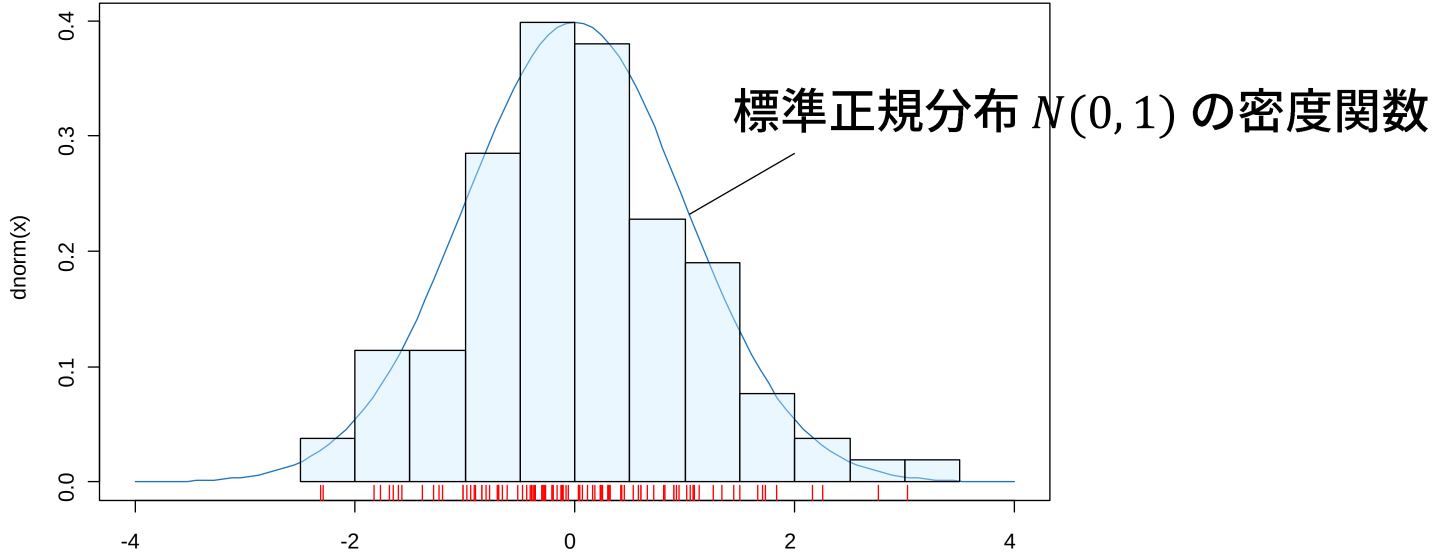


- 数直線上にプロットしてみる
  - グラフ > ドットプロット
- 分布形状を推定してみる
  - グラフ > ヒストグラム
  - （参考）グラフ > 密度推定

ノンパラメトリック密度推定: 関数の形が未知の場合にデータから密度関数を推定

標準正規分布に従うデータ

- 標準正規分布に従うデータは平均0の周りに密集
  - $P(-3 \leq X \leq 3) = 0.9974$ なので-3から3を外れる点はめったに出ない



（参考: 上記の図のRスクリプト）

```
curve(dnorm, -4, 4, xlab="")
rug(NormalSamples$obs, xlim=c(-4, 4), xlab="")
par(new=T)
hist(NormalSamples$obs, xlim=c(-4, 4), xlab="", xaxt="n", ylab="", yaxt="n", main="")
```

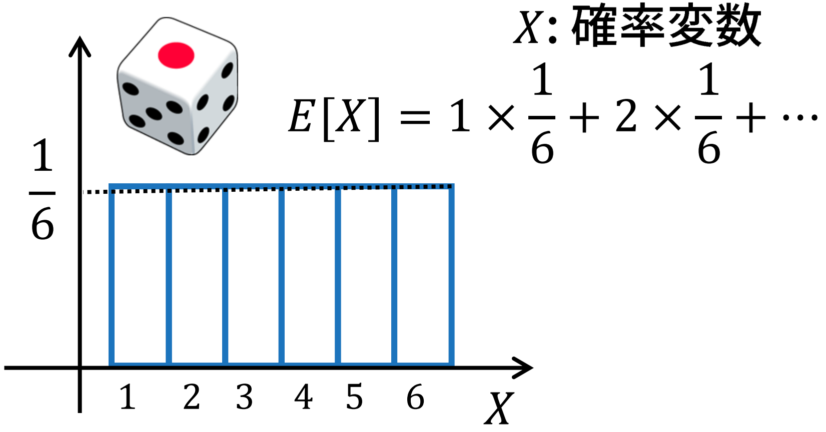
本日の講義内容

- テキスト
  - 「統計学入門」4～5章
- 確率，確率変数，確率分布
  - [Rcmdr] 確率分布に従う乱数を生成してみる
- 期待値の定義，線形性，平均と分散の加法性
  - 次回よく使う
- 上側（下側，両側）確率，パーセント点
  - [Rcmdr] 正規分布の上側確率やパーセント点を表示してみる
- 宿題（兼レポート予告）

期待値の計算方法

- 期待値は「確率変数」に対して定義される
- 離散分布の期待値

$$E[X] = \sum_k x_k \, P(X = x_k)$$



- 連続分布の期待値

$$E[X] = \int_{-\infty}^{\infty} x \, f_X(x) dx$$

$f_X(x)$は確率密度関数 $\int_{-\infty}^{\infty} f_X(x) dx = 1$

教科書によっては$E[X]$を$E(X)$と書くことも

記述統計の平均値との比較

- 記述統計におけるデータの平均値

1が3回，2が4回，3が2回，4が1回

$$\bar{x} = \frac{3+1+2+3+2+2+2+1+4+1}{10}$$

  - 度数を使って書き直してみる

$$\bar{x} = \frac{1\times 3 \; + \; 2\times 4 \; + \; 3\times 2 \; + \; 4\times 1}{3+4+2+1} = 1\cdot 0.3 + 2\cdot 0.4 + 3\cdot 0.2 + 4\cdot 0.1$$

  - $\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_K f_K}{N} = x_1 \frac{f_1}{N} + x_2 \frac{f_2}{N} + \cdots + x_K \frac{f_K}{N} = \boxed{\sum_{k=1}^{K} x_k p_k}$
    - ただし$N = f_1 + f_2 + \cdots + f_K$，$p_k = \frac{f_k}{N}$とおいた
- 一方，確率変数Xに対しては具体的な度数f_kではなく分布$P(X = x_k)$が与えられる：期待値$\boxed{E[X] = \sum_k x_k P(X = x_k)}$
- 期待値$E[X]$を確率変数Xの平均と呼ぶことも多い

確率変数の平均と分散

- 確率変数Xの平均 (mean)
  - $\mu = E[X]$
    （ギリシャ文字μはアルファベットmと対応（大文字はどちらもM））

平均からのバラつき

- 確率変数Xの分散 (variance)
  - $\sigma^2 = V[X] = E[(X - E[X])^2] = E[(X - \mu)^2]$

- 確率変数Xの標準偏差 (standard deviation)
  - $\sigma = \sqrt{V[X]}$
    （ギリシャ文字σはアルファベットsと対応）

平均と分散の加法性（重要！）

- 期待値と分散
  - $E[aX+b]=aE[X]+b$ （期待値の線形性）　　p.96-, 148-
  - $V[aX+b]=a^2V[X]$
- 二つの確率変数 X,Y について以下が成立
  - 平均に関して $E[X+Y]=E[X]+E[Y],\quad E[X-Y]=E[X]-E[Y]$
  - 分散に関して $V[X\pm Y]=V[X]+V[Y]$ （X,Y が無相関のとき成立）

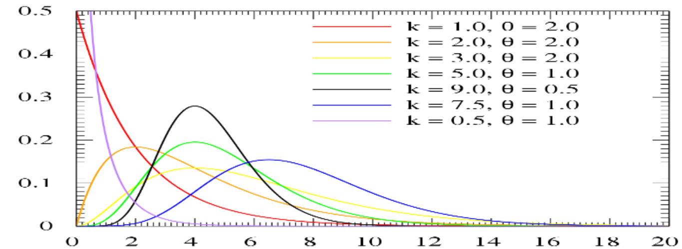
独立ならば無相関

- 確率変数 $X_1,X_2,\dots,X_n$ について以下が成立
  - 平均に関して $E[X_1+X_2+\cdots+X_n]=E[X_1]+E[X_2]+\cdots+E[X_n]$
  - 分散に関して $V[X_1+X_2+\cdots+X_n]=V[X_1]+V[X_2]+\cdots+V[X_n]$
    - ただし分散の加法性は、確率変数が互いに無相関のときに成立（互いに独立でも成立）

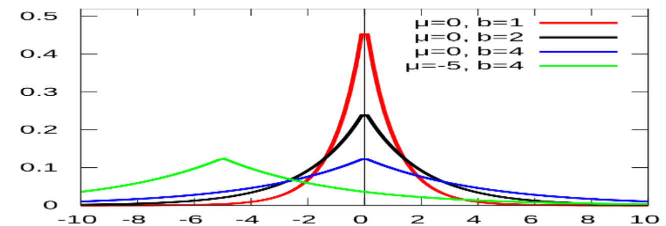
独立ならば無相関

歪度と尖度

- 平均 (mean)　$\mu = E[X]$
- 分散 (variance)　$\sigma^2 = E[(X-\mu)^2]$
- 歪度 (skewness)　$\dfrac{E\left[(X-\mu)^3\right]}{\sigma^3}$
  - 分布の非対称性を表す．正規分布は0
  - <u>正ならば裾が右に長い</u>
- 尖度 (kurtosis)　$\dfrac{E\left[(X-\mu)^4\right]}{\sigma^4} - 3$
  - 分布の尖り具合を表す．正規分布は0 （-3を付けない定義もある）
  - 正ならば中心付近が尖っているだけでなく<u>裾が両側に長い（重い）</u>
- モーメント（積率）(moment)　$\mu_r = E[(X-\mu)^r]$
  - 平均のまわりのr次モーメント．原点まわりで定義することもある



ガンマ分布(Wikipediaより)



ラプラス分布(Wikipediaより)

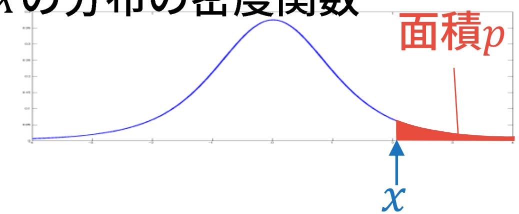
本日の講義内容

- テキスト
  - 「統計学入門」4～5章
- 確率，確率変数，確率分布
  - [Rcmdr] 確率分布に従う乱数を生成してみる
- 期待値の定義，線形性，平均と分散の加法性
  - 次回よく使う
- 上側（下側，両側）確率，パーセント点
  - [Rcmdr] 正規分布の上側確率やパーセント点を表示してみる
- 宿題（兼レポート予告）

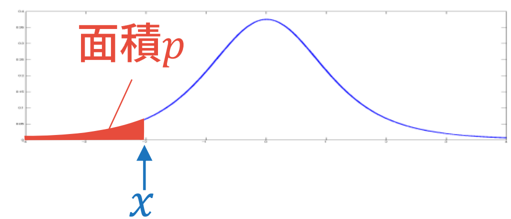
上側・下側・両側確率

- ## 上側確率
  - Xが値x以上となる確率: $p = P(X \geq x)$
  - 片側検定におけるp値

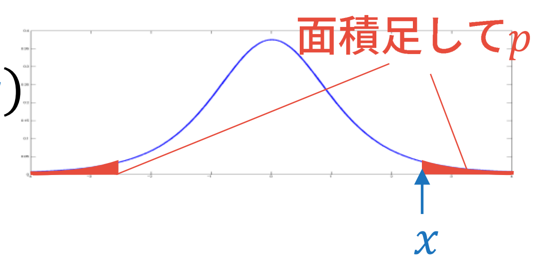
Xの分布の密度関数



- ## 下側確率
  - Xが値x以下となる確率: $p = P(X \leq x)$
  - 片側検定におけるp値



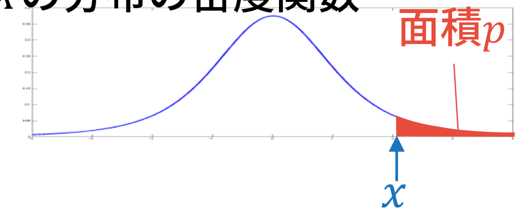
- ## 両側確率（分布は左右対称とする）
  - Xが値$-x$以下かx以上となる確率: $p = P(|X| \geq x)$
    - $p = P(X \leq -x) + P(X \geq x) = 2P(X \geq x)$
  - 両側検定におけるp値



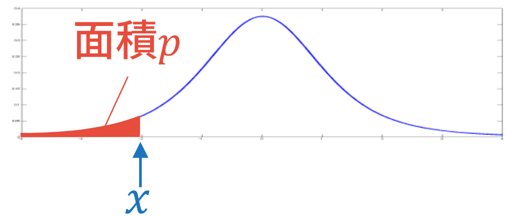
パーセント点 (percentile)

- **上側$100p$ %点**（上側確率$100p$ %点ともいう）
  - $P(X \geq x) = p$ を満たすx　（例: $p = 0.05$ で上側5%点）
  - 上側確率(密度関数の右側面積)が p になる値x

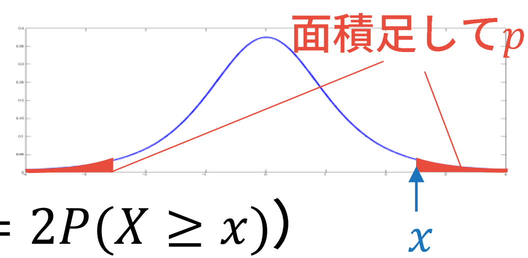
Xの分布の密度関数



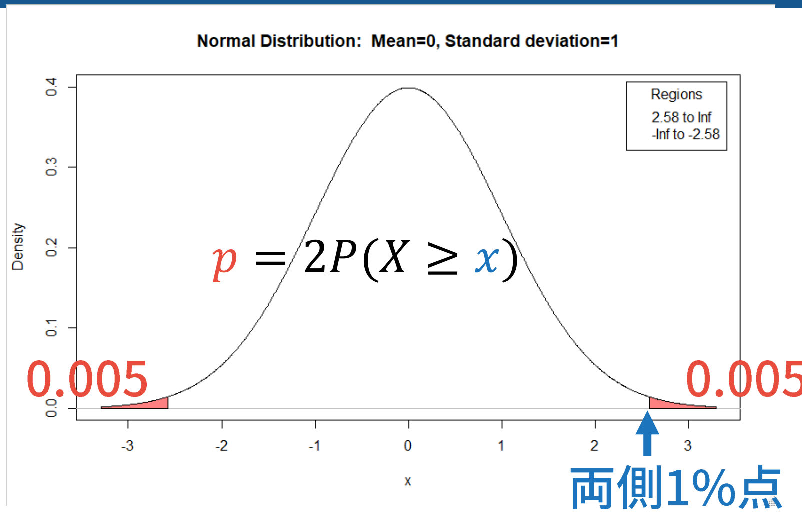
- **下側$100p$ %点**（単に$100p$ %点といえばこれ）
  - $P(X \leq x) = p$ を満たすx
  - $p = \frac{1}{4}, \frac{2}{4}, \frac{3}{4}$ は四分位点

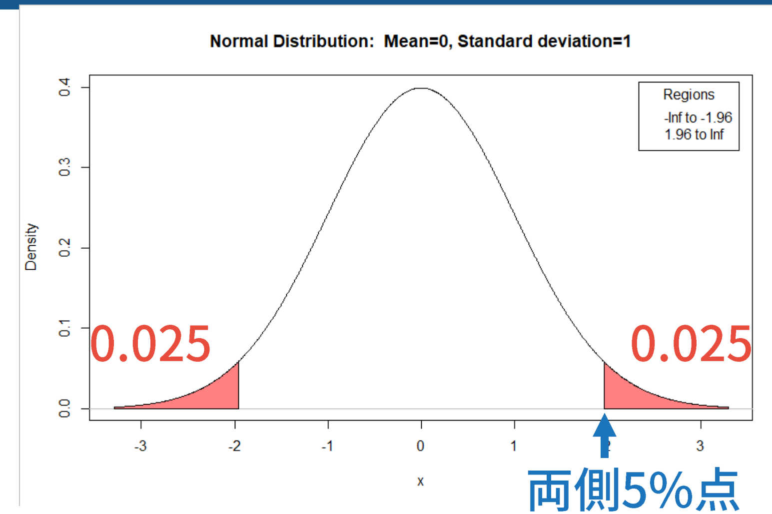


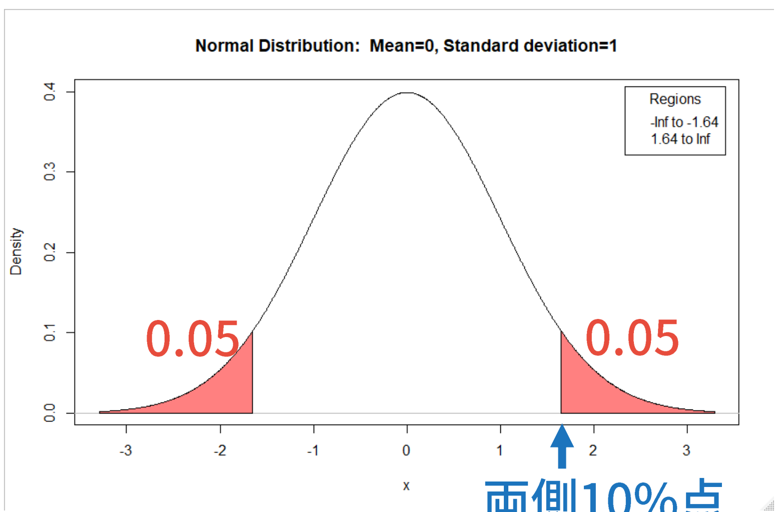
- **両側$100p$ %点**（分布は左右対称とする）
  - $P(|X| \geq x) = p$ を満たすx
    - $P(X \geq x) = p/2$ を満たすx（$\because P(|X| \geq x) = 2P(X \geq x)$）
  - 両側確率(密度関数の左右面積の和)が p になる値x

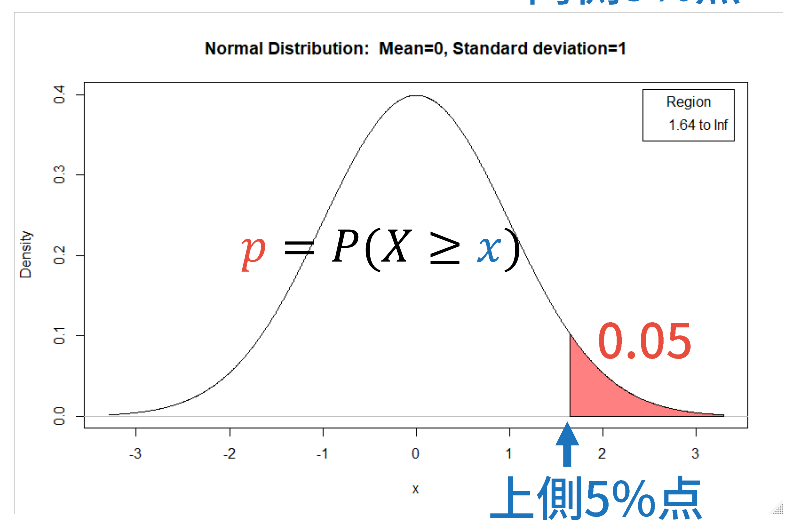


標準正規分布のパーセント点



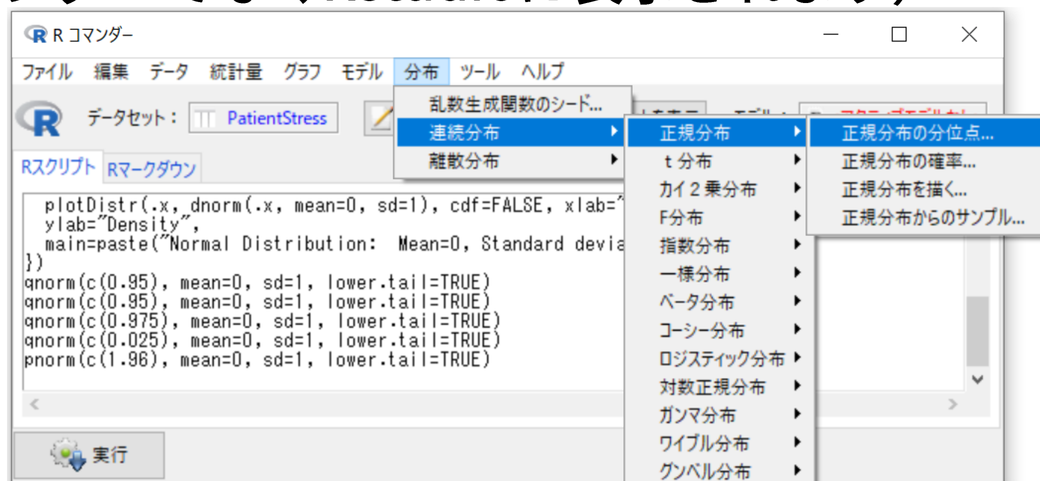






[Rcmdr] パーセント点と上側・下側確率

- 正規分布のパーセント点
  - 分布 > 連続分布 > 正規分布 > 正規分布の分位点
  - 0.975 や 0.025 で確認（Rコマンダーでなく Rstudioに表示されます）
- 正規分布の上側（下側）確率
  - 分布 > 連続分布 > 正規分布 > 正規分布の確率
  - 1.96 や -1.96 で確認（Rコマンダーでなく Rstudioに表示されます）
  - p 値の計算にも利用できる
- t分布でも試してみる



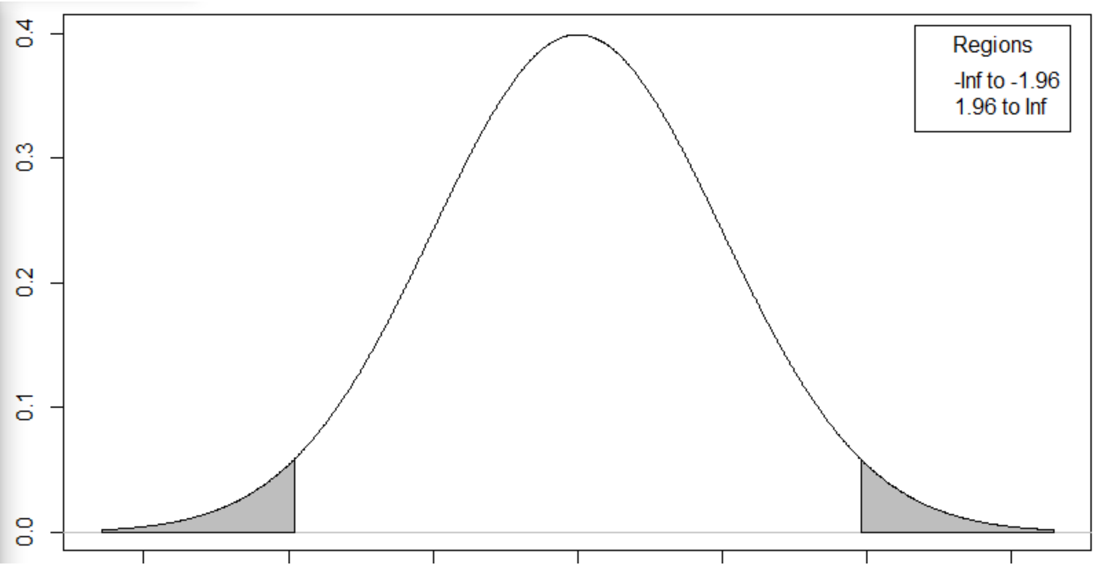
宿題（兼レポート予告）

1. 身長のデータ（先週の height.csv）を標準化せよ．
   - データ > アクティブデータセット内の変数の管理 > 変数の標準化
   - 標準化後の「Z.身長」のヒストグラムを表示し，横軸の値を確認せよ
2. 標準正規分布を描き，以下の領域を色付けせよ．
   - 両側確率5%の領域
   - 上側確率5%の領域
3. 標準正規分布の次の値を小数点以下2桁で示し (Rコマンダー or Excel)，テキストp.280の「付表1 正規分布表」と照らし合わせて確認せよ．
   - 両側1パーセント点，5パーセント点，10パーセント点
   - 上側1パーセント点，5パーセント点，10パーセント点

[Rcmdr] 正規分布を描き領域を色付けする

- 分布 > 連続分布 > 正規分布 > 正規分布を描く
  - 密度関数のプロットに表示する領域の指定: 分位点
  - 色付けする領域: （下側確率で指定する）
    - 両側確率5%を描くなら:

      領域1: from 0 to 0.025, 領域2: from 0.975 to 1（色も適当に指定)
      （パーセント点 1.96も凡例からわかる）

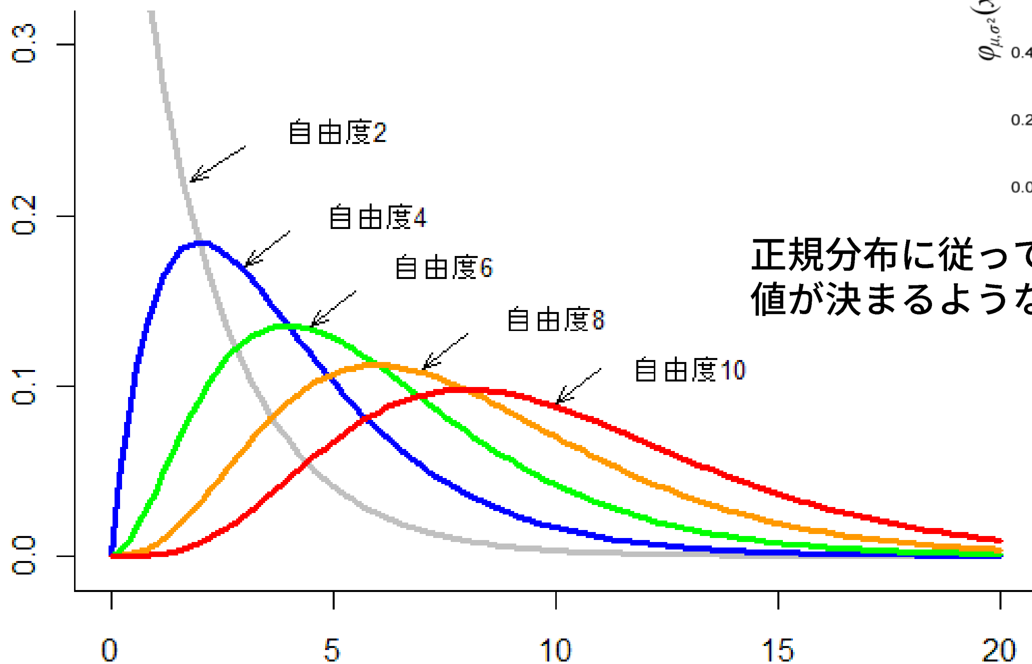


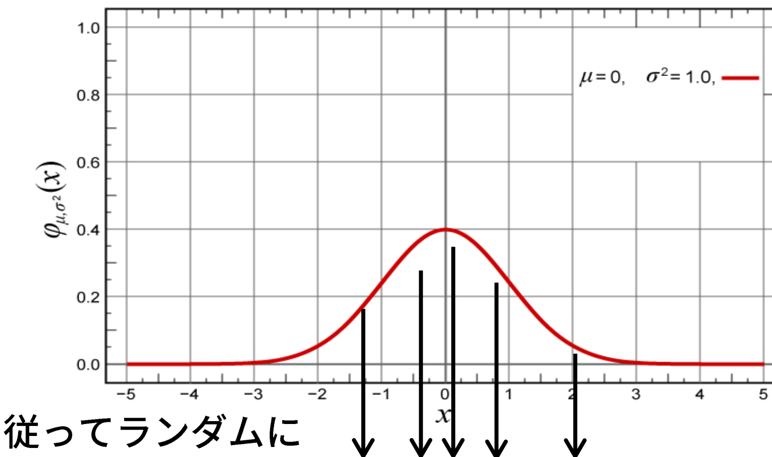
発展（後日出てくる分布）

- 余力があれば是非これ以降の演習をRコマンダーで行ってください
  - 「確率・統計」の後半で出てくる分布です

χ^2分布

- 標準正規分布に従う複数の確率変数の二乗和が従う確率分布
  - k個の和ならば自由度k





正規分布に従ってランダムに
値が決まるようなk個の変数

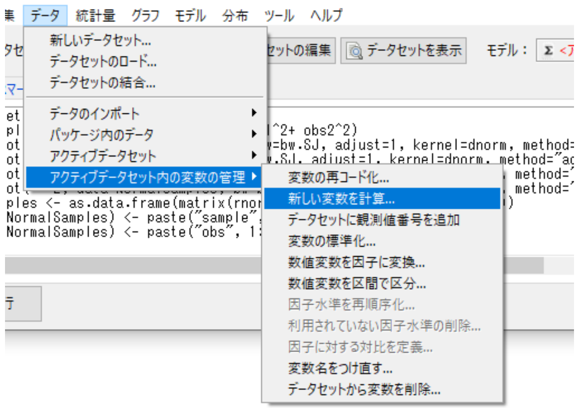
$$\underbrace{X_1, X_2, X_3, \cdots, X_k}$$

$$Z = \sum_{i=1}^{k} X_i^2$$

二乗和Zは自由度kのχ^2分布に

[Rcmdr] 正規分布からχ^2分布を生成

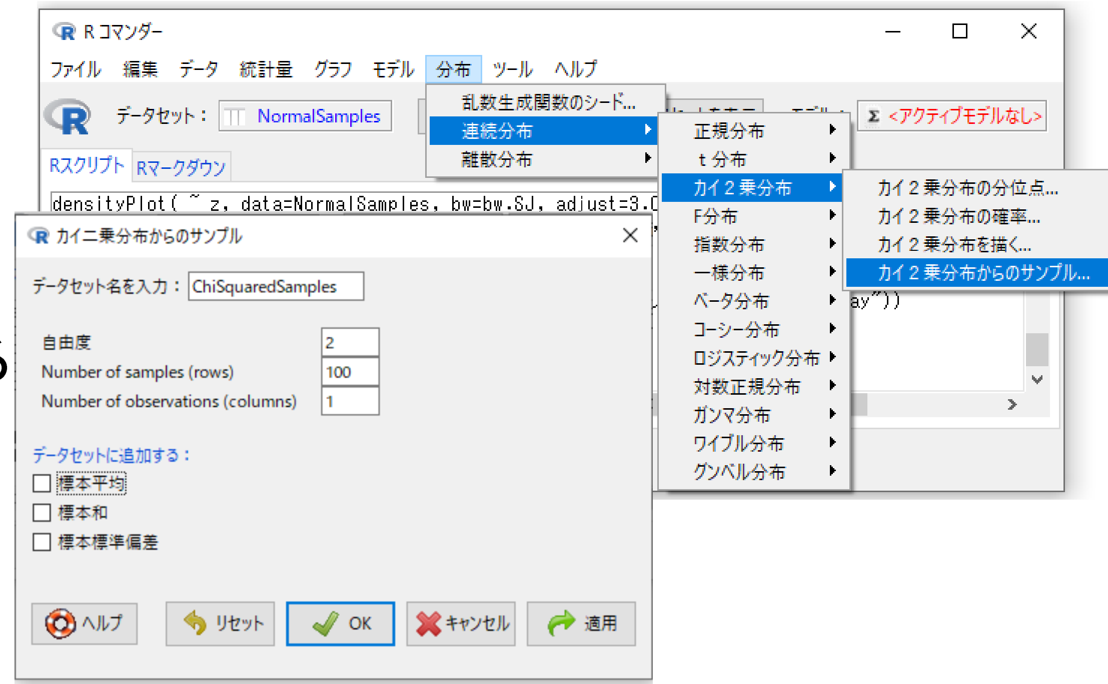
- 標準正規分布に従う乱数を生成
  - 分布 > 連続分布 > 正規分布 > 正規分布からのサンプル
  - 平均: 0，標準偏差: 1
  - Number of samples: 100
  - Number of observations: 2（自由度kに相当）
  - データセットは上書き
  - 「データセットを表示」で確認



- 新たな列を作成
  - データ > アクティブデータセット内の変数の管理 > 新しい変数を計算
  - 新しい変数名: z，計算式: obs1^2 + obs2^2
  - 「データセットを表示」で列が追加されたことを確認
- z の分布を確認
  - グラフ > ドットプロット（ドットチャート，ヒストグラム，箱ひげ）

[Rcmdr] 直接χ^2分布に従うデータを生成

- 元から用意されている分布の機能を利用
  - （分布 > 連続分布 > カイ2乗分布 > カイ2乗分布を描く）
  - 分布 > 連続分布 > カイ2乗分布 > カイ2乗分布からのサンプル
    - 自由度: 2
    - … (rows): 100
    - … (columns): 1
  - グラフ > ヒストグラム
  - 自由度3, 4でも試してみる



その他のよく用いられる分布（いくつかは後日）

- 離散分布
  - ポアソン分布：$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}\quad(\lambda>0),\ \ k=0,1,2,\ldots$
  - 幾何分布：$P(X=k)=p(1-p)^{k-1}\quad(0<p<1),\ k=1,2,3,\ldots$
  - 超幾何分布：$P(X=k)=\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}=\frac{\binom{n}{k}\binom{N-n}{K-k}}{\binom{N}{K}}$
- 連続分布
  - 一様分布：密度関数 $f(x)=\begin{cases}\frac{1}{\beta-\alpha} & (\alpha\leq x\leq\beta)\\ 0 & otherwise\end{cases}$
  - 指数分布：密度関数 $f(x)=\begin{cases}\lambda e^{-\lambda x} & (x\geq 0)\\ 0 & otherwise\end{cases}\quad(\lambda>0)$
  - ガンマ分布：密度関数 $f(x)=\frac{1}{\Gamma(k)\theta^k}x^{k-1}e^{-\frac{x}{\theta}}\ \ (x>0)\ \ (k>0,\theta>0)$