Excel 演習で理解する統計学入門

第1章　データの種類

1.1　質的データと量的データ

例題1-1　質的データと量的データのちがいを確認する

次のデータについて，質的データの書式を「文字列」に，量的データの書式を「数値」に設定しましょう．

	A	B	C	D	E	F	G	H	I
1	学生番号	入学年度	学部	年齢	出身地（県）	数学の点数	数学の評価	数学検定（級）	満足度
2	287010	2028	経営	20	愛知	70	3	3	5
3	296111	2029	経済	19	愛知	55	2	3	3
4	297131	2029	経営	20	鹿児島	60	3	3	5
5	288024	2028	法	25	愛知	100	5	準2	2
6	288033	2028	法	21	高知	45	2	4	8
7	307142	2030	経営	19	福島	60	3	3	10
8	297027	2029	経営	19	三重	75	3	3	5
9	276100	2027	経済	22	沖縄	70	3	4	5
10	278080	2027	法	21	愛知	30	1	5	1
11	307066	2030	経営	18	鹿児島	85	4	3	7

10人の学生についての情報

例題 1-1　質的データと量的データのちがいを確認する

データは,「計算できるかどうか」という視点で，質的データと量的データに分けられます．

Point! **質的データと量的データ**

質的データは分類をあらわし，量的データは数量をあらわします．

例題 1-1　質的データと量的データのちがいを確認する

さらに，質的データは,「順序に意味があるか」といった観点から，名義尺度と順序尺度に分けられます.

量的データは,「比に意味があるか」といった観点から，間隔尺度と比例尺度に分けられます.

例題 1-1　質的データと量的データのちがいを確認する

質的データ

- 「同一性」のみをあらわしている質的データは名義尺度とよばれます.

学生番号や電話番号は，数値であらわされていても，その数値の大きさに意味はなく，計算することにも意味はありません．いわば「名前」のようなものなので，質的データです．出身地や性別は，分類するためのものであり，質的データということになります.

以上のような，同じかちがうかのみに意味がある質的データは名義尺度とよばれ,「同一性」のみをあらわしています.

例題 1-1　質的データと量的データのちがいを確認する

- 「同一性」「順序性」をあらわしている質的データは順序尺度とよばれます.

段階評価，検定の級，順位のように,「同一性」のみではなく「順序性」もあらわす質的データもあります．年代，時間帯などの幅のある区間はこの例になります.
これらにおいては，順序（大小関係）には意味がありますが，間隔（差）にも比率にも意味がありません.

このような質的データは順序尺度とよばれます.

例題1-1　質的データと量的データのちがいを確認する

量的データ

- 「同一性」「順序性」「等間隔性」のみをあらわしている量的データは間隔尺度とよばれます．

和暦，温度（摂氏・華氏）などは，数値であらわされ，その数値に意味があり，差にも意味があるものです．これらは量的データと解釈できます．原点（0）には絶対的な意味がなく，比率には意味がありませんが，間隔（差）には意味があります．

このような量的データは間隔尺度とよばれ,「同一性」「順序性」だけではなく,「等間隔性」もあらわしています．

例題1-1　質的データと量的データのちがいを確認する

- 「同一性」「順序性」「等間隔性」「等比性」をあらわしている量的データは比例尺度とよばれます.

年齢，身長，速度，値段，睡眠時間のように，原点（0）に絶対的な意味があり，間隔（差）にも比率にも意味がある量的データもあります.

これらは，比例尺度とよばれ,「同一性」「順序性」「等間隔性」だけではなく「等比性」もあらわしています.

例題 1-1　質的データと量的データのちがいを確認する

データの種類	尺度	意味				例
		区別	順序	差	比	
質的データ	名義尺度	○	×	×	×	学生番号，電話番号，出身地，性別
	順序尺度	○	○	×	×	段階評価，級，順位，年代，時間帯
量的データ	間隔尺度	○	○	○	×	和暦，温度（摂氏・華氏），偏差値
	比例尺度	○	○	○	○	年齢，身長，速度，値段，睡眠時間

例題 1-1　質的データと量的データのちがいを確認する

	A	B	C	D	E	F	G	H	I
1	学生番号	入学年度	学部	年齢	出身地（県）	数学の点数	数学の評価	数学検定（級）	満足度
2	287010	2028	経営	20	愛知	70	3	3	5
3	296111	2029	経済	19	愛知	55	2	3	3
4	297131	2029	経営	20	鹿児島	60	3	3	5
5	288024	2028	法	25	愛知	100	5	準2	2
6	288033	2028	法	21	高知	45	2	4	8
7	307142	2030	経営	19	福島	60	3	3	10
8	297027	2029	経営	19	三重	75	3	3	5
9	276100	2027	経済	22	沖縄	70	3	4	5
10	278080	2027	法	21	愛知	30	1	5	1
11	307066	2030	経営	18	鹿児島	85	4	3	7

学生番号，学部，出身地，数学の評価，数学検定，満足度は質的データと解釈することができ，入学年度，年齢，数学の点数は量的データと解釈することができます．

1.2　データセットの種類

例題 1-2　データセットの種類を確認する

次の 6 セットのデータセットは，ある店舗の同じ商品についての情報であり，それぞれはクロスセクションデータです．これらを，1 セットの時系列データとしてまとめましょう．

	A	B	C
1		価格（円）	売上個数
2	1日目	1200	720

	A	B	C
1		価格（円）	売上個数
2	2日目	1080	762

	A	B	C
1		価格（円）	売上個数
2	3日目	1260	600

	A	B	C
1		価格（円）	売上個数
2	4日目	1080	732

	A	B	C
1		価格（円）	売上個数
2	5日目	1200	666

	A	B	C
1		価格（円）	売上個数
2	6日目	960	858

それぞれはクロスセクションデータ（売上個数についての情報）

例題 1-2　データセットの種類を確認する

一定の形式に整えられたデータの集まりのことをデータセットといいます．

Excel で分析するデータセットは行と列で構成されています．ここで，行は横方向の並びのことで，列は縦方向の並びのことです．

例題 1-2　データセットの種類を確認する

Excel でおこなうデータ分析では，「知りたい対象」が行に入り，その対象についての「知りたい情報」が列に入る形が基本です．

	A	B	C	D	E	F	G	H	I
1	学生番号	入学年度	学部	年齢	出身地（県）	数学の点数	数学の評価	数学検定（級）	満足度
2	287010	2028	経営	20	愛知	70	3	3	5
3	296111	2029	経済	19	愛知	55	2	3	3
4	297131	2029	経営	20	鹿児島	60	3	3	5
5	288024	2028	法	25	愛知	100	5	準2	2
6	288033	2028	法	21	高知	45	2	4	8
7	307142	2030	経営	19	福島	60	3	3	10
8	297027	2029	経営	19	三重	75	3	3	5
9	276100	2027	経済	22	沖縄	70	3	4	5
10	278080	2027	法	21	愛知	30	1	5	1
11	307066	2030	経営	18	鹿児島	85	4	3	7

10 人の学生についての情報

例題 1-2　データセットの種類を確認する

このとき，変数とは，その列に「どんな情報が入っているか」をあらわすものです．たとえば「出身地」という変数なら，愛知，鹿児島，高知，$\cdots$ などのデータが入ります．

この例では,「学生番号」「入学年度」「学部」「年齢」「出身地」「数学の点数」「数学の評価」「数学検定」「満足度」という変数があります．

また，対象は 10 人の学生ということになります．

例題 1-2　データセットの種類を確認する

「時間に関係するかどうか」という視点で，データセットの種類を確認しましょう．

Point! **クロスセクションデータ，時系列データ，パネルデータ**

同じ時点における複数の対象についてのデータはクロスセクションデータ（横断面データ）とよばれます．

また，対象が 1 つであっても，同じ時点における複数の変数についてのデータであれば，クロスセクションデータに分類されます．

例題 1-2　データセットの種類を確認する

Point!

クロスセクションデータ，時系列データ，パネルデータ

一方，特定の 1 つの対象について，同じ変数を，複数の時点にわたって継続的に観測したデータは時系列データとよばれます．

そして，特定の複数の対象について，同じ変数を，複数の時点にわたって継続的に観測したデータはパネルデータとよばれます．

例題 1-2　データセットの種類を確認する

次のように 1 セットの時系列データとしてまとめることができます．

	A	B	C
1		価格（円）	売上個数
2	1日目	1200	720
3	2日目	1080	762
4	3日目	1260	600
5	4日目	1080	732
6	5日目	1200	666
7	6日目	960	858
8			

時系列データ（売上個数についての情報）

1.3　母集団と標本

例題 1-3　母集団と標本のちがいを確認する

次のデータを母集団として，そこから 50 名分のデータを無作為に標本として抽出してみましょう．

	A	B	C	D	E	F	G	H	I	J
1	母集団									
2	学生番号	学生所属	年次	点数						
3	30-6001	経済学部	1	42						
4	30-6002	経済学部	1	42						
5	30-6003	経済学部	1	42						
6	30-6004	経済学部	1	80						
7	30-6005	経済学部	1	10						
8	30-6006	経済学部	1	80						
9	30-6007	経済学部	1	66						
10	30-6008	経済学部	1	62						
11	30-6009	経済学部	1	90						
12	30-6010	経済学部	1	62						
13	30-6011	経済学部	1	52						
14	30-6012	経済学部	1	50						
15	30-6013	経済学部	1	54						
16	30-6014	経済学部	1	58						
17	30-6015	経済学部	1	80						

Sheet1

学生についての情報

例題 1-3　母集団と標本のちがいを確認する

「調査対象すべてかどうか」という視点で，データの種類を確認しましょう．

Point! **母集団と標本**

調査対象のデータ全体は母集団とよばれます．

また，母集団の情報を推測するために母集団から抽出された一部分は標本（サンプル）とよばれます．ここで，標本はデータの集まりであり，標本に含まれるデータの個数を標本サイズ（標本の大きさ）といいます．

例題 1-3　母集団と標本のちがいを確認する

Point!

全数調査と標本調査

母集団（データ全体）を調べることは全数調査といわれます．

一方，標本（一部分）を調べて母集団について推測することは標本調査といわれます．

無作為に標本を抽出することを無作為抽出（ランダムサンプリング）といいます．

例題 1-3　母集団と標本のちがいを確認する

	A	B	C	D	E	F	G	H	I	J
1	母集団						標本1			
2	学生番号	学生所属	年次	点数	乱数		学生番号	学生所属	年次	点数
3	27-8063	法学部	4	44	0.60787		27-8063	法学部	4	44
4	29-7002	経営学部	2	20	0.173555		29-7002	経営学部	2	20
5	28-6009	経済学部	3	90	0.316509		28-6009	経済学部	3	90
6	28-7110	経営学部	3	70	0.610803		28-7110	経営学部	3	70
7	27-7102	経営学部	4	44	0.178268		27-7102	経営学部	4	44
8	29-7069	経営学部	2	50	0.621448		29-7069	経営学部	2	50
9	30-7062	経営学部	1	90	0.614709		30-7062	経営学部	1	90
10	30-7122	経営学部	1	40	0.003599		30-7122	経営学部	1	40
11	30-9504	人間生活部	1	54	0.491282		30-9504	人間生活部	1	54
12	27-6077	経済学部	4	46	0.435177		27-6077	経済学部	4	46
13	29-6060	経済学部	2	38	0.558209		29-6060	経済学部	2	38
14	29-6115	経済学部	2	38	0.340324		29-6115	経済学部	2	38
15	30-6065	経済学部	1	90	0.076846		30-6065	経済学部	1	90
16	30-9013	人間生活部	1	50	0.181345		30-9013	人間生活部	1	50
17	27-7106	経営学部	4	40	0.420134		27-7106	経営学部	4	40

Sheet1　+

例題 1-3　母集団と標本のちがいを確認する

Point!

記述統計学と推測統計学

データに対して，平均値などの統計量を求めたり，グラフで可視化したりなど，そのデータの特徴をより明確に把握する方法を研究する統計学の分野は，記述統計学といわれます．

一方,「一部分を調べて全体を知る」ための方法を研究する，標本調査にもとづく統計学の分野は，推測統計学といわれます．

第2章　データの可視化

2.1　質的データの可視化

例題 2-1　質的データを集計する

女性の人数，男性の人数をそれぞれ求めてみましょう．

	A	B	C	D
1	番号	性別	車のボディタイプ	住所
2	1	男性	セダン	名古屋市
3	2	男性	ミニバン	春日井市
4	3	女性	ミニバン	名古屋市
5	4	男性	ハッチバック	清須市
6	5	男性	ミニバン	長久手市
7	6	男性	セダン	名古屋市
8	7	男性	ステーションワゴン	犬山市
9	8	男性	ミニバン	名古屋市
10	9	女性	ミニバン	春日井市
11	10	男性	ステーションワゴン	春日井市
12	11	男性	ミニバン	あま市
13	12	女性	SUV	名古屋市
14	13	女性	SUV	春日井市
15	14	男性	ミニバン	あま市
16	15	男性	ミニバン	名古屋市
17	16	男性	SUV	春日井市

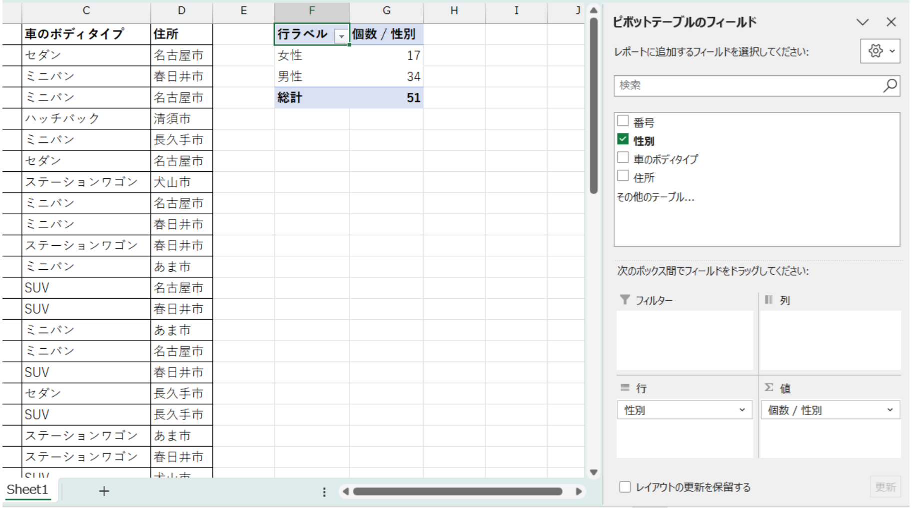
Sheet1

ある店舗における利用者の所有車についての情報

例題 2-1　質的データを集計する

「性別」のような質的データでは,「男性」「女性」それぞれの人数をかぞえたり，全体に占める割合を計算したり，円グラフなどで可視化することで，データの傾向を把握できることがあります.

例題 2-1　質的データを集計する



車のボディタイプ	住所
セダン	名古屋市
ミニバン	春日井市
ミニバン	名古屋市
ハッチバック	清須市
ミニバン	長久手市
セダン	名古屋市
ステーションワゴン	犬山市
ミニバン	名古屋市
ミニバン	春日井市
ステーションワゴン	春日井市
ミニバン	あま市
SUV	名古屋市
SUV	春日井市
ミニバン	あま市
ミニバン	名古屋市
SUV	春日井市
セダン	長久手市
SUV	長久手市
ステーションワゴン	あま市
ステーションワゴン	春日井市

行ラベル	個数 / 性別
女性	17
男性	34
総計	**51**

例題 2-2　質的データを円グラフにする

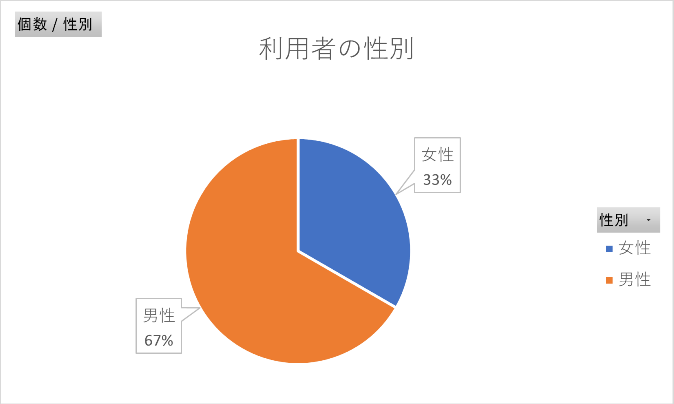
例題 2-1 のデータについて，円グラフを作成してみましょう．ここで，男性の割合，女性の割合がそれぞれ表示されるようにしましょう．

	A	B	C	D
1	番号	性別	車のボディタイプ	住所
2	1	男性	セダン	名古屋市
3	2	男性	ミニバン	春日井市
4	3	女性	ミニバン	名古屋市
5	4	男性	ハッチバック	清須市
6	5	男性	ミニバン	長久手市
7	6	男性	セダン	名古屋市
8	7	男性	ステーションワゴン	犬山市
9	8	男性	ミニバン	名古屋市
10	9	女性	ミニバン	春日井市
11	10	男性	ステーションワゴン	春日井市
12	11	男性	ミニバン	あま市
13	12	女性	SUV	名古屋市
14	13	女性	SUV	春日井市
15	14	男性	ミニバン	あま市
16	15	男性	ミニバン	名古屋市
17	16	男性	SUV	春日井市

Sheet1 +

ある店舗における利用者の所有車についての情報

例題 2-2　質的データを円グラフにする



例題 2-2　質的データを円グラフにする

Point!
円グラフ

円グラフでは，各扇形の中心角の大きさとその扇形があらわす数値の大きさが比例します．質的データにおいては，円全体を 100%として各項目の割合を扇形であらわします．割合どうしを比較したり，全体と一部を比較したりするのに適しています．

ここで，質的データにおいて，変数がとる値の種類を項目とよびます．たとえば，「性別」をあらわす変数では，「男性」「女性」が項目にあたります．

例題2-3　質的データを棒グラフにする

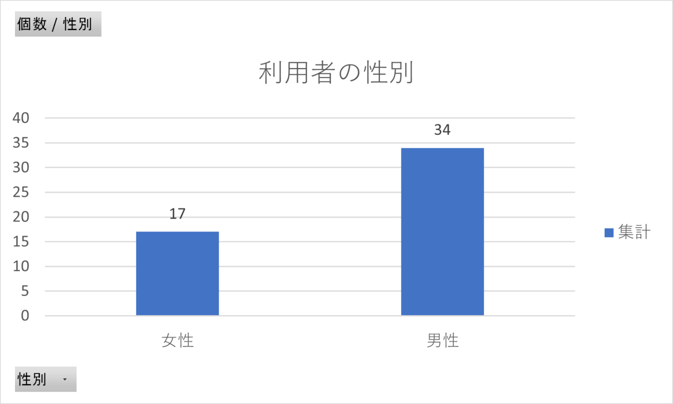
例題2-1のデータについて，棒グラフを作成してみましょう．ここで，男性の人数，女性の人数がそれぞれ表示されるようにしましょう．

番号	性別	車のボディタイプ	住所
1	男性	セダン	名古屋市
2	男性	ミニバン	春日井市
3	女性	ミニバン	名古屋市
4	男性	ハッチバック	清須市
5	男性	ミニバン	長久手市
6	男性	セダン	名古屋市
7	男性	ステーションワゴン	犬山市
8	男性	ミニバン	名古屋市
9	女性	ミニバン	春日井市
10	男性	ステーションワゴン	春日井市
11	男性	ミニバン	あま市
12	女性	SUV	名古屋市
13	女性	SUV	春日井市
14	男性	ミニバン	あま市
15	男性	ミニバン	名古屋市
16	男性	SUV	春日井市

Sheet1

ある店舗における利用者の所有車についての情報

例題 2-3　質的データを棒グラフにする



例題2-3　質的データを棒グラフにする

Point!

棒グラフ

棒グラフは棒の長さで数値の大きさをあらわします．質的データにおいては，各項目に属するデータの個数や割合を比較するのに適しています．

例題 2-4　質的データをクロス集計する

例題 2-1 のデータについて，男女別に，車の各ボディタイプを所有しているそれぞれの人数を求めてみましょう．また，求めた結果を使い，100% 積み上げ横棒グラフを作成してみましょう．

番号	性別	車のボディタイプ	住所
1	男性	セダン	名古屋市
2	男性	ミニバン	春日井市
3	女性	ミニバン	名古屋市
4	男性	ハッチバック	清須市
5	男性	ミニバン	長久手市
6	男性	セダン	名古屋市
7	男性	ステーションワゴン	犬山市
8	男性	ミニバン	名古屋市
9	女性	ミニバン	春日井市
10	男性	ステーションワゴン	春日井市
11	男性	ミニバン	あま市
12	女性	SUV	名古屋市
13	女性	SUV	春日井市
14	男性	ミニバン	あま市
15	男性	ミニバン	名古屋市
16	男性	SUV	春日井市

Sheet1

ある店舗における利用者の所有車についての情報

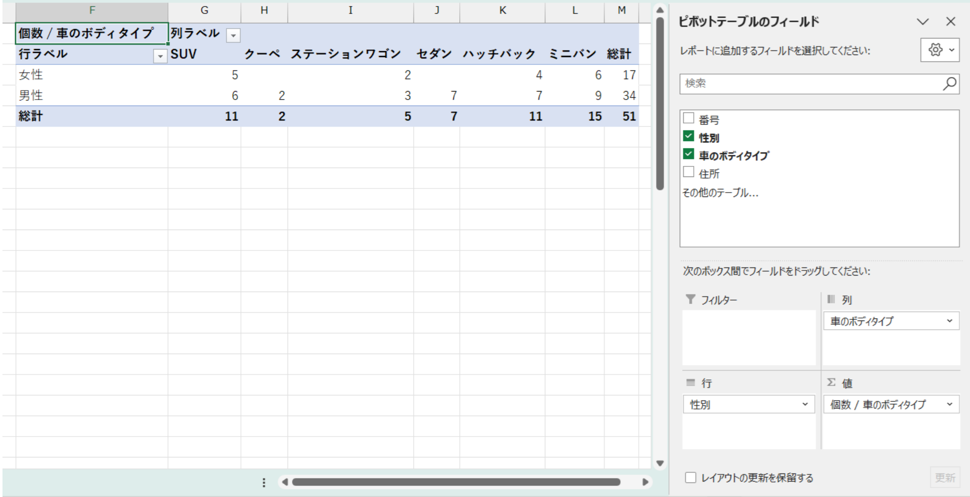
例題 2-4　質的データをクロス集計する

Point!
クロス集計

クロス集計とは，質的データにおいて，複数の変数を組み合わせて集計する方法です．行と列に異なる変数についての項目を配置し，それぞれの組み合わせに該当するデータの個数や割合を表にまとめます．

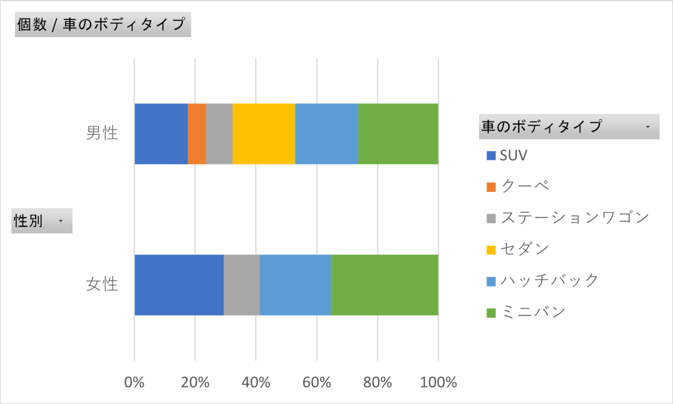
この例題でおこなうのは,「性別 × 車のボディタイプ」のクロス集計であり，男女別に，車の各ボディタイプ（セダン，クーペ，．．．）を所有しているそれぞれの人数を一覧で確認します．

例題2-4　質的データをクロス集計する



個数 / 車のボディタイプ	列ラベル						
行ラベル	SUV	クーペ	ステーションワゴン	セダン	ハッチバック	ミニバン	総計
女性	5		2		4	6	17
男性	6	2	3	7	7	9	34
総計	11	2	5	7	11	15	51

例題 2-4　質的データをクロス集計する



例題2-4　質的データをクロス集計する

Point!
100% 積み上げ棒グラフ

100% 積み上げ棒グラフでは，棒の長さとその棒があらわす数値の大きさが比例します．質的データにおいては，棒全体の長さを100%として各項目の割合を棒の長さであらわします．割合どうしを比較したり，全体と一部を比較したりするのに適しています．

2.2　量的データの可視化

例題 2-5　量的データをヒストグラムにする

ヒストグラムを作成してみましょう．区間の幅は 500 にします．

	A	B
1	日付	来客数
2	7月19日	4342
3	7月20日	5390
4	7月21日	1439
5	7月22日	1491
6	7月23日	1479
7	7月24日	1579
8	7月25日	1952
9	7月26日	3949
10	7月27日	4856
11	7月28日	1304
12	7月29日	2195
13	7月30日	1593
14	7月31日	1835
15	8月1日	976
16	8月2日	5198
17	8月3日	4810

Sheet1

ある市営プールにおける来客数の情報

例題 2-5　量的データをヒストグラムにする

「来客数」のような量的データについては，質的データのように,「同じデータ（同じ数値）」の出現回数を集計したり，割合を求めたりする方法では全体像を把握できないことがあります．

そこで，データの分布の様子を見るために，量的データを一定幅の区間ごとに分けて質的データにし，それぞれの区間に入っているデータの個数（度数）をまとめます．

これを度数分布表とよび, これをグラフにしたものをヒストグラムとよびます．

例題 2-5　量的データをヒストグラムにする

なお，データの分布とは，データがどの範囲や項目にどのくらい存在しているかを示すものです．それぞれの区間における度数により示すものだけでなく，質的データの各項目に属するデータの個数により示すものも含みます．

例題2-5　量的データをヒストグラムにする

Point! **ヒストグラム**

量的データを一定幅の区間ごとに分けて，その区間に入っているデータの個数を調べると，データの分布の様子をみることができます．

この一定幅の区間のことを階級といい，この各階級に含まれるデータの個数のことを度数といいます．各階級の度数が全度数に占める割合は相対度数とよばれます．

例題 2-5　量的データをヒストグラムにする

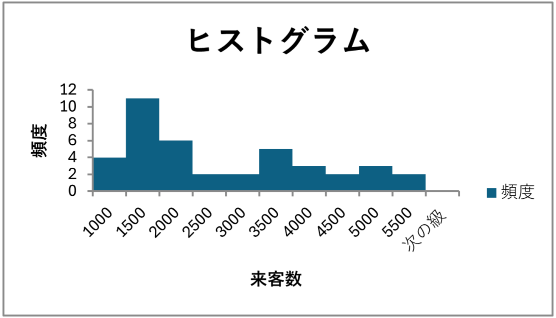
Point!

ヒストグラム

そして，小さい階級から順に度数を加えていったものは累積度数，小さい階級から順に相対度数を加えていったものは累積相対度数とよばれます．

また，階級ごとの度数を棒の高さで表現したグラフはヒストグラムとよばれます．

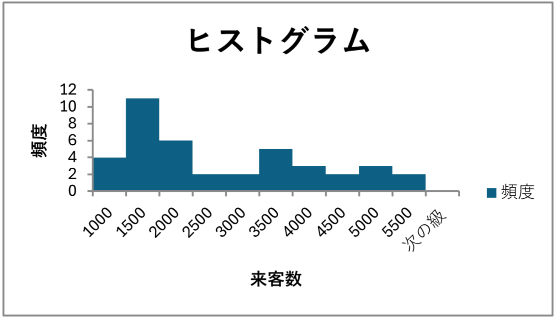
例題 2-5　量的データをヒストグラムにする



ヒストグラムの各棒の高さは，左からそれぞれ,「1000 以下」,「1000 より大きく 1500 以下」,「1500 より大きく 2000 以下」，· · · ,「5000 より大きく 5500 以下」の区間の度数をあらわしています.

例題 2-6　量的データの度数分布表を作成する

例題 2-5 で作成したヒストグラムの階級（データ区間）における，相対度数，累積度数，累積相対度数も求めてみましょう．



例題 2-6　量的データの度数分布表を作成する

	D	E	F	G	H	I	J	K	L	M	N
	最大値	5390		データ区間		データ区間	頻度	相対度数	累積度数	累積相対度数	
	最小値	830		1000		1000	4	0.1	4	0.1	
				1500		1500	11	0.275	15	0.375	
				2000		2000	6	0.15	21	0.525	
				2500		2500	2	0.05	23	0.575	
				3000		3000	2	0.05	25	0.625	
				3500		3500	5	0.125	30	0.75	
				4000		4000	3	0.075	33	0.825	
				4500		4500	2	0.05	35	0.875	
				5000		5000	3	0.075	38	0.95	
				5500		5500	2	0.05	40	1	
						次の級	0				
						合計	40				

一番大きい階級の累積度数は合計と等しくなり，一番大きい階級の累積相対度数は 1 になることが確認できます．

例題 2-7　時系列データを折れ線グラフにする

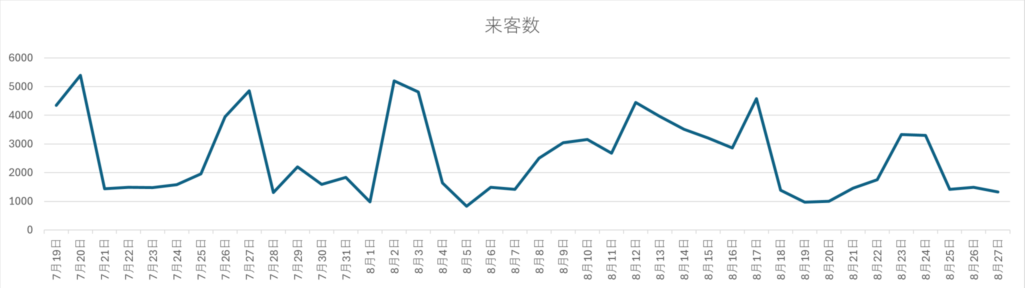
日付を横軸として，来客数の折れ線グラフを作成してみましょう．

	A	B
1	日付	来客数
2	7月19日	4342
3	7月20日	5390
4	7月21日	1439
5	7月22日	1491
6	7月23日	1479
7	7月24日	1579
8	7月25日	1952
9	7月26日	3949
10	7月27日	4856
11	7月28日	1304
12	7月29日	2195
13	7月30日	1593
14	7月31日	1835
15	8月1日	976
16	8月2日	5198
17	8月3日	4810

Sheet1

ある市営プールにおける来客数の情報

例題 2-7　時系列データを折れ線グラフにする



例題 2-7　時系列データを折れ線グラフにする

Point!

折れ線グラフ

折れ線グラフは点を線分で結ぶので，データの推移をあらわすのに適しています．時系列データについて作成すると，時間による傾向がつかみやすくなります．

例題 2-8　量的データと量的データを散布図にする

散布図を作成してみましょう.

	A	B	C
1	番号	身長（cm）	体重（kg）
2	1	163	54
3	2	170	65
4	3	180	66
5	4	175	75
6	5	161	49
7	6	178	80
8	7	175	61
9	8	167	95
10	9	182	101
11	10	152	43
12	11	155	50

11 人の身長と体重についての情報

例題2-8　量的データと量的データを散布図にする

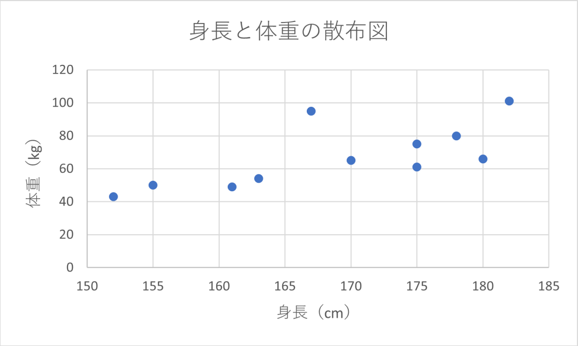
Point!

散布図

2 つの変数についてのデータの組を xy 平面上に点としてあらわしたグラフは散布図とよばれます．散布図を作成することにより，2 変数の関係が見やすくなったり，外れ値（大きく外れた値）を発見しやすくなったりすることがあります．

例題 2-8　量的データと量的データを散布図にする

2 つの変数がそれぞれ原因と結果をあらわしている可能性があるならば，原因と思われる変数を x（横軸），結果と思われる変数を y（縦軸）にして散布図を作成しましょう．



第3章　データの中心とデータのばらつき

3.1　代表値

例題 3-1　平均値を求める

各生徒における 5 科目の平均点を求めてみましょう．また，数学と理科の 2 科目の平均点も求めてみましょう．

	A	B	C	D	E	F	G
1	**科目**	**生徒A**	**生徒B**	**生徒C**	**生徒D**	**生徒E**	**生徒F**
2	国語	84	89	81	47	90	56
3	数学	41	78	70	78	70	87
4	英語	50	60	74	97	80	94
5	理科	52	80	70	70	53	57
6	社会	85	79	70	54	91	81

5 科目についての生徒 A から生徒 F のテストの点数

例題 3-1　平均値を求める

たとえば生徒 A の 5 科目の平均点は

$$\frac{84+41+50+52+85}{5} = \frac{312}{5} = 62.4$$

より，62.4 点となります．

このような,「データの合計をデータの個数で割って求める平均値」を相加平均値といいます．

例題 3-1　平均値を求める

外れ値（大きく外れた値）はデータの合計に影響を与えるので，相加平均値は外れ値の影響を受けやすいという特徴があります．

Point!

相加平均値

$$
\text{相加平均値} = \frac{\text{合計}}{\text{個数}}
$$

例題 3-1　平均値を求める

これより，

$$\text{合計} = \text{相加平均値} \times \text{個数}$$

となり，合計は，相加平均値にデータの個数をかけると求められます．つまり，相加平均値をデータの個数分足すと合計になるということになります：

$$\text{相加平均値} + \text{相加平均値} + \cdots + \text{相加平均値} = \text{合計}$$

例題 3-1　平均値を求める

相加平均値のように，データの中心的傾向を 1 つの数値で代表したものは代表値とよばれます．代表値はデータを代表する値ということであり，平均値，中央値，そして，最頻値などが使われます．

なお，平均値にはいろいろな算出方法（定義式の種類）がありますが，以下，単に平均値と記すときは相加平均値のことを指すこととします．

例題 3-1　平均値を求める

Point!

相加平均値の意味

「もしどの点数も相加平均値 62.4 であると強引にみなしてしまったとしても，それらの合計はもともとの合計 312 と変わらない」というものが相加平均値です．

$84 + 41 + 50 + 52 + 85 = 312$　（もともとの合計）

$62.4 + 62.4 + 62.4 + 62.4 + 62.4 = 312$　（どの点数も 62.4 としたときの合計）

例題 3-1　平均値を求める

	A	B	C	D	E	F	G	H
1	**科目**	**生徒A**	**生徒B**	**生徒C**	**生徒D**	**生徒E**	**生徒F**	
2	国語	84	89	81	47	90	56	
3	数学	41	78	70	78	70	87	
4	英語	50	60	74	97	80	94	
5	理科	52	80	70	70	53	57	
6	社会	85	79	70	54	91	81	
7	5科目の平均点	62.4	77.2	73	69.2	76.8	75	
8	数学と理科の平均点	46.5	79	70	74	61.5	72	
9								

例題 3-2　中央値，最頻値を求める

各科目の点数における平均値，中央値，最頻値をそれぞれ求めてみましょう．

	A	B	C	D	E	F
1	生徒	国語	数学	英語	理科	社会
2	A	84	41	50	52	85
3	B	89	78	60	80	79
4	C	81	70	74	70	70
5	D	47	78	97	70	54
6	E	90	70	80	53	91
7	F	56	87	94	57	81

生徒 A から生徒 F についての 5 科目のテストの点数

例題 3-2　中央値，最頻値を求める

Point! **中央値**

中央値とは，データを大きさの順に並べ替えたときの真ん中に位置する値のことをいいます．

- データの個数が奇数のとき（つまり真ん中の値が 1 つのとき）は，その真ん中の値をそのまま中央値とします．
- データの個数が偶数のとき（つまり真ん中の値が 2 つのとき）は,「真ん中の 2 つの値の平均値」を中央値とします．

例題 3-2　中央値，最頻値を求める

Point! **最頻値**

最も頻繁に現れるデータのことを最頻値とよびます．

例題 3-2　中央値，最頻値を求める

平均値は外れ値（大きく外れた値）の影響を受けやすいという特徴がありますが，中央値は真ん中の 1 つまたは 2 つの値のみから決まり，それ以外の値の影響は受けません．

中央値は外れ値の影響を受けにくいといえます．

例題3-2　中央値，最頻値を求める

そして，最頻値はもっともよく現れるデータ「そのもの」です．それゆえ，中央値と同様，外れ値の影響を受けにくいということになります．

また，最頻値だけを知っても，それ以外のデータがどのようなものかはわからないので，最頻値はデータ全体の傾向を把握するのには適さない場合があることに注意しましょう．

例題 3-2　中央値，最頻値を求める

	A	B	C	D	E	F	G
1	生徒	国語	数学	英語	理科	社会	
2	A	84	41	50	52	85	
3	B	89	78	60	80	79	
4	C	81	70	74	70	70	
5	D	47	78	97	70	54	
6	E	90	70	80	53	91	
7	F	56	87	94	57	81	
8	平均点	74.5	70.66667	75.83333	63.66667	76.66667	
9	中央値	82.5	74	77	63.5	80	
10	最頻値	#N/A	78	#N/A	70	#N/A	
11			70				

「#N/A」は該当なしという意味で，この場合，最頻値が存在しないことを示しています．数学の点数の最頻値は 2 つあるので，スピルにより 2 つの値 78，70 の両方が表示されていることが確認できます．

例題 3-3　平均値，中央値，最頻値を求めて比較する

下記のアルバイトの時給についてのデータ（単位：円）の平均値，中央値，最頻値を求めましょう．

$$1300,\ 1800,\ 1600,\ 1500,\ 1700,\ 1700,\ 1300,\ 1300,\ 17000,\ 1300$$

例題 3-3　平均値，中央値，最頻値を求めて比較する

	A	B
1		1300
2		1800
3		1600
4		1500
5		1700
6		1700
7		1300
8		1300
9		17000
10		1300
11		
12	**平均値**	3050
13	**中央値**	1550
14	**最頻値**	1300
15		

アルバイトの時給についてのデータ（単位：円）

例題 3-3　平均値，中央値，最頻値を求めて比較する

平均値は 3050 円，中央値は 1550 円，最頻値は 1300 円と計算されました．

平均値だけ，かなり大きくなってしまいました．これは，極端に大きいデータ「17000 円」のせいで合計が大きくなってしまうからです．平均値だけ見ると，実態より高く見えます．

例題 3-3　平均値，中央値，最頻値を求めて比較する

一方，中央値は 5 番目に大きい値「1500 円」と 6 番目に大きい値「1600 円」のみから決まり，それ以外の値の影響は受けていません．データの中心的傾向を反映しているといえます．

また，最頻値「1300 円」はもっともよく現れるデータ「そのもの」であり，もっともよく見られる時給をあらわしています．

3.2　分散，標準偏差

例題 3-4　各データから平均値を引いたものを調べる

身長，体重それぞれについて，平均値を求めて偏差を計算し，偏差の棒グラフを作成しましょう．

	A	B	C	D	E	F	G
1	番号	身長（cm）	体重（kg）	身長の偏差（cm）	体重の偏差（kg）	身長の偏差の2乗（cm^2）	体重の偏差の2乗（kg^2）
2	1	177	89				
3	2	165	71				
4	3	153	53				
5	4	177	99				
6	5	180	95				
7	6	159	59				
8	7	156	45				
9	8	174	89				
10	9	168	75				
11	10	171	53				
12	平均値						
13	分散						
14	標準偏差						

10 人の身長と体重についての情報

例題 3-4　各データから平均値を引いたものを調べる

Point!

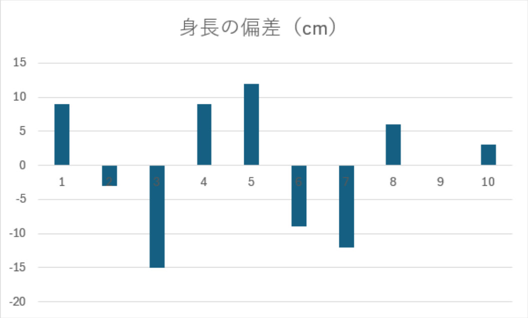
偏差

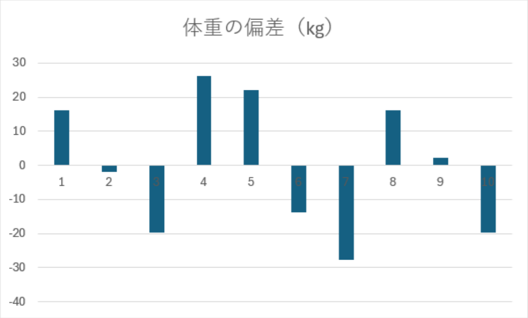
各データから平均値を引いたものを偏差といいます．

例題 3-4　各データから平均値を引いたものを調べる

	A	B	C	D	E	F	G
1	番号	身長（cm）	体重（kg）	身長の偏差（cm）	体重の偏差（kg）	身長の偏差の2乗（cm^2）	体重の偏差の2乗（kg^2）
2	1	177	89	9	16.2		
3	2	165	71	-3	-1.8		
4	3	153	53	-15	-19.8		
5	4	177	99	9	26.2		
6	5	180	95	12	22.2		
7	6	159	59	-9	-13.8		
8	7	156	45	-12	-27.8		
9	8	174	89	6	16.2		
10	9	168	75	0	2.2		
11	10	171	53	3	-19.8		
12	平均値	168	72.8				
13	分散						
14	標準偏差						

例題 3-4　各データから平均値を引いたものを調べる





例題 3-5　データのばらつきの大きさを調べる

例題 3-4 において，身長，体重それぞれについて，偏差（平均値との差）の棒グラフを作成しました．

それを見ると，身長は平均値を中心としておおよそプラスマイナス「10 cm 程度」の範囲でばらついていて，体重は平均値を中心としておおよそプラスマイナス「20 kg 程度」の範囲でばらついていることが確認できます．

例題 3-5　データのばらつきの大きさを調べる

この「10 cm 程度」,「20 kg 程度」といった「ばらつきの大きさ」の判断は主観的なものですが，これをどのように計算して求めればいいのでしょうか．

「標準的な偏差（標準的な平均値との差）」はどのように算出されるのかを考えてみましょう．

例題 3-5　データのばらつきの大きさを調べる

偏差は 0 を中心にプラスマイナスにばらついているため，偏差の平均値はいつも 0 になってしまいます．

例題 3-5　データのばらつきの大きさを調べる

偏差の 2 乗の平均値については，データのばらつきが大きくなるほど大きくなり，データのばらつきが小さくなくほど小さくなる数値であるといえます．
つまり，これは,「データのばらつきの大きさをあらわす数値」だといえるでしょう．

このような値のことを分散とよびます．

例題 3-5　データのばらつきの大きさを調べる

	A	B	C	D	E	F	G	H
1	番号	身長（cm）	体重（kg）	身長の偏差（cm）	体重の偏差（kg）	身長の偏差の2乗（cm^2）	体重の偏差の2乗（kg^2）	
2	1	177	89	9	16.2	81	262.44	
3	2	165	71	-3	-1.8	9	3.24	
4	3	153	53	-15	-19.8	225	392.04	
5	4	177	99	9	26.2	81	686.44	
6	5	180	95	12	22.2	144	492.84	
7	6	159	59	-9	-13.8	81	190.44	
8	7	156	45	-12	-27.8	144	772.84	
9	8	174	89	6	16.2	36	262.44	
10	9	168	75	0	2.2	0	4.84	
11	10	171	53	3	-19.8	9	392.04	
12	平均値	168	72.8	0	0	**81**	**345.96**	
13	分散							
14	標準偏差							
15								

例題 3-5　データのばらつきの大きさを調べる

Point! **分散**

偏差の 2 乗の平均値，つまり，偏差の 2 乗和をデータの個数で割ったものを分散とよびます．

身長の分散は $81\ \mathrm{cm}^2$，体重の分散は $345.96\ \mathrm{kg}^2$ となります．

例題 3-5　データのばらつきの大きさを調べる

身長は平均値を中心としておおよそプラスマイナス「10 cm 程度」の範囲で，体重は平均値を中心としておおよそプラスマイナス「20 kg 程度」の範囲でばらついています．

しかし，身長の分散は 81 cm^2，体重の分散は 345.96 kg^2 であり，「標準的な偏差（標準的な平均値との差）」とするには値が大きすぎます．しかも，偏差の単位は「cm」，「kg」ですが，分散の単位は「cm^2」，「kg^2」になってしまいます．

例題 3-5　データのばらつきの大きさを調べる

そこで，たとえば，身長の分散 $81\ \mathrm{cm}^2$ の正の平方根（$\sqrt{\ }$）をとってみると，

$$\sqrt{\mathbf{81\,cm^2}} = \mathbf{9\,cm}$$

となります．

例題 3-5　データのばらつきの大きさを調べる

こうすると，単位も元のデータと同じ「cm」に戻り，さらに,「標準的な偏差（標準的な平均値との差）」といっていい値になりました．

このような値のことを標準偏差とよびます．

例題 3-5　データのばらつきの大きさを調べる

Point! **標準偏差**

$$標準偏差 = \sqrt{分散}$$

例題 3-5　データのばらつきの大きさを調べる

	A	B	C	D	E	F	G
1	番号	身長（cm）	体重（kg）	身長の偏差（cm）	体重の偏差（kg）	身長の偏差の2乗（cm^2）	体重の偏差の2乗（kg^2）
2	1	177	89	9	16.2	81	262.44
3	2	165	71	-3	-1.8	9	3.24
4	3	153	53	-15	-19.8	225	392.04
5	4	177	99	9	26.2	81	686.44
6	5	180	95	12	22.2	144	492.84
7	6	159	59	-9	-13.8	81	190.44
8	7	156	45	-12	-27.8	144	772.84
9	8	174	89	6	16.2	36	262.44
10	9	168	75	0	2.2	0	4.84
11	10	171	53	3	-19.8	9	392.04
12	平均値	168	72.8	0	0	**81**	**345.96**
13	分散	**81**	**345.96**				
14	標準偏差	9	18.6				

身長の標準偏差は 9 cm，体重の標準偏差は 18.6 kg となります．

例題3-6　平均値に対するデータのばらつきの大きさを調べる

Kくんと体重と飼い猫Mちゃんの体重それぞれについて，平均値と比べてのばらつきの大きさを調べてみましょう．

	A	B	C
1	測定した月	Kくんの体重（kg）	Mちゃんの体重（kg）
2	1月	58.6	2.8
3	2月	59.1	2.7
4	3月	58.9	2.7
5	4月	58.6	2.6
6	5月	58.1	2.7
7	6月	57.5	2.5
8	7月	56.9	2.4
9	8月	55.2	2.5
10	9月	56.3	2.5
11	10月	56.3	2.6
12	11月	55.4	2.7
13	12月	58.9	2.9
14			

Kくんの体重と飼い猫Mちゃんの体重についての情報

例題 3-6　平均値に対するデータのばらつきの大きさを調べる

6	5月	58.1	2.7
7	6月	57.5	2.5
8	7月	56.9	2.4
9	8月	55.2	2.5
10	9月	56.3	2.5
11	10月	56.3	2.6
12	11月	55.4	2.7
13	12月	58.9	2.9
14	平均値	57.48333333	2.633333333
15	標準偏差	1.360044934	0.137436854
16			
17			

平均値は，K くんの体重については約 57.5 kg，M ちゃんの体重については約 2.6 kg になりました．
また,「データのばらつきの大きさ」をあらわす数値である標準偏差は，K くんの体重については約 1.36 kg，M ちゃんの体重については約 0.14 kg になりました．

例題 3-6　平均値に対するデータのばらつきの大きさを調べる

K くんの体重のほうが標準偏差が大きいことがわかりました．

こんどは，平均値と比べての標準偏差の大きさを比較してみましょう．
つまり，標準偏差を平均値で割って，変動係数を求めてみましょう．

例題 3-6　平均値に対するデータのばらつきの大きさを調べる

変動係数は単位をもたない数値であり，データのばらつきの大きさを平均値に対する割合で示します．そのため，単位が異なるデータや平均値が異なるデータのばらつきの大きさを比較する際に用いることができます．

標準偏差が「絶対的な揺れ幅」をあらわすのに対し，変動係数は「平均値に対しての相対的な揺れ幅」をあらわします．

例題3-6　平均値に対するデータのばらつきの大きさを調べる

Point! **変動係数**

平均値が0でないとき，標準偏差を平均値で割ったものを変動係数といいます．

例題 3-6　平均値に対するデータのばらつきの大きさを調べる

7	6月	57.5	2.5
8	7月	56.9	2.4
9	8月	55.2	2.5
10	9月	56.3	2.5
11	10月	56.3	2.6
12	11月	55.4	2.7
13	12月	58.9	2.9
14	平均値	57.48333333	2.633333333
15	標準偏差	1.360044934	0.137436854
16	変動係数	0.023659813	0.05219121
17			

K くんの体重の変動係数は約 0.024，M ちゃんの体重の変動係数は約 0.052 であり，変動係数は M ちゃんの体重のほうが大きいことがわかりました．

例題 3-6　平均値に対するデータのばらつきの大きさを調べる

これより，K くんの体重は標準偏差が大きいものの，平均値に対してはＭちゃんの体重のほうが相対的にばらつきが大きいことが確認できます．

第 4 章　2 項分布

4.1　ベルヌーイ試行

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

サイコロを 60 回振ったときの 1 の目が出る割合を計算してみましょう．そして，この結果がサイコロを振ったときに 1 の目が出る確率に近い値になることを確かめましょう．

同様に，1 以外の目が出る割合も計算し，この結果が 1 以外の目が出る確率に近い値になることも確かめましょう．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

また，サイコロを 60 回振ったときの 1 の目が出る回数が 10 である確率を計算しましょう．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

試行と事象

「サイコロを振る」というような，結果が偶然に起こるような実験や観測は試行といわれ，試行の結果のいくつかからなる集合のことを事象といわれます．

とくに，試行によって起こりうる結果全体の集合のことは全事象とよばれます．

また，元が 1 つの集合であらわされる事象は根元事象とよばれます．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

事象は一般に，集合の形であらわされます．たとえば,「サイコロを振る」という試行についての全事象 U は

$$U = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$$

とあらわすことができます．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

そして,「1 の目が出る」という事象 A は,

$$A = \{1\}$$

というように U の部分集合であらわすことができます．これは，元が 1 つの集合であらわされる事象なので，根元事象です．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

またたとえば,「奇数の目が出る」という事象 O は,

$$O = \{1, 3, 5\}$$

というようにあらわすことができます.

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

サイコロを振ったときに $\mathbf{1}$ の目が出る確率というのは，全事象

$$\boldsymbol{U} = \{\mathbf{1},\ \mathbf{2},\ \mathbf{3},\ \mathbf{4},\ \mathbf{5},\ \mathbf{6}\}$$

の $\mathbf{6}$ 通りのなかでの 1 の目が出る割合を考えているのです．

1 の目が出る場合は $A = \{1\}$ で 1 通りだけなので，サイコロを振ったときに 1 の目が出る確率は $1/6$ であることがわかります．つまり，次のように計算しています．

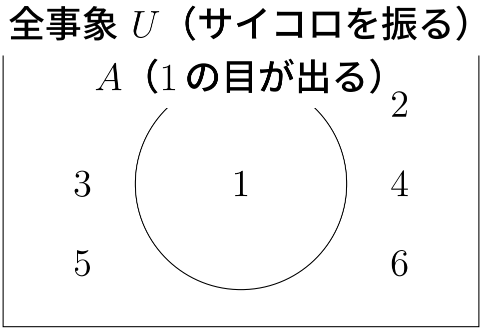
例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

$$1\text{の目が出る確率} = \frac{\text{事象}\,A\,\text{の元の個数}}{\text{全事象}\,U\,\text{の元の個数}}$$

$$= \frac{1\text{の目が出る場合の数}}{\text{出る目についてのすべての場合の数}}$$

$$= \frac{1}{6}$$

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる



サイコロを振ったときに 1 の目が出る確率をあらわすベン図

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

サイコロを振ったときに出る目のように,「同じ程度に起こることが期待できる」という意味は同様に確からしいと表現されることがあります.

一般に，全事象 U の各根元事象が起こることが互いに同様に確からしいとすると，事象 A が起こる確率は次のように求められます.

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

Point!

確率の求め方

$$\text{事象 } A \text{ が起こる確率} = \frac{\text{事象 } \boldsymbol{A} \text{ の元の個数}}{\text{全事象 } \boldsymbol{U} \text{ の元の個数}}$$

$$= \frac{\text{事象 } A \text{ が起こる場合の数}}{\text{起こりうるすべての場合の数}}$$

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

「1 以外の目が出る」ことは,「1 の目が出る」ことの余事象になり, 次が成り立ちます.

$$「1\text{以外の目が出る確率}」= 1 - 「1\text{の目が出る確率}」= 1 - \frac{1}{6} = \frac{5}{6}$$

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

一般に，事象 A に対して，A が起こらないという事象を A の余事象といい，$\overline{A}$ であらわします．このとき，次が成り立ちます．

Point!

余事象の確率

「$\overline{A}$ が起こる確率」$= 1 -$「A が起こる確率」

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

	A	B	C	D	E	F	G	H
1	試行		「1」が出る回数	「1」が出る割合	←実験	理論→	「1」が出る確率	
2	1		10	0.166666667			0.166666667	
3	6		「1以外」が出る回数	「1以外」が出る割合			「1以外」が出る確率	
4	1		50	0.833333333			0.833333333	
5	3							
6	3							
7	1							
8	6							
9	6							
10	3							
11	3							
12	3							
13	5							
14	2							
15	1							
16	6							
17	3							

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

はじめの 10 回は「1」が出て残りの 50 回は「1 以外」が出る確率は

$$\text{「1 回目に「1」が出る確率」} \times \cdots \times \text{「10 回目に「1」が出る確率」}$$

$$\times \text{「11 回目に「1 以外」が出る確率」} \times \cdots \times \text{「60 回目に「1 以外」が出る確率」}$$

$$= \left(\frac{1}{6}\right)^{10} \times \left(\frac{5}{6}\right)^{50}$$

と計算して求められます（独立事象の乗法定理）．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

独立事象の乗法定理

事象 A，B が互いに独立であるとは，互いに影響を与えないということであり，次が成り立つことをいいます．

$$A\text{ かつ }B\text{ が起こる確率} = A\text{ が起こる確率} \times B\text{ が起こる確率}$$

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

「1」が出る回数が 10 であるときの「1」と「1 以外」の並び方のパターンは，上記の

　はじめの 10 回は「1」が出て残りの 50 回は「1 以外」が出る

以外にもあります．全部で何通りあるのでしょうか．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

これは，60 回のなかから「1」が出る回を 10 個選ぶときの選び方の総数であり，

相異なる 60 個のものから 10 個を重複なく選んだものの総数
（組み合わせの総数）

となります．つまり，これは ${}_{60}\mathrm{C}_{10}$ とあらわされます．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

これら ${}_{60}\mathrm{C}_{10}$ 通りのパターンそれぞれについて，それが起こる確率は上記で求めた

はじめの 10 回は「1」が出て残りの 50 回は「1 以外」が出る確率

$$\left(\frac{1}{6}\right)^{10} \times \left(\frac{5}{6}\right)^{50}$$

と等しくなります．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

Point! **組み合わせ**

一般に，n 個のものから k 個を重複なく選んだものを組み合わせといい，その総数は ${}_n\mathrm{C}_k$ とあらわされます．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

たとえば，${}_5\mathrm{C}_3$ は次のようになります．

$$
{}_5\mathrm{C}_3 = \frac{{}_5\mathrm{P}_3}{3!} = \frac{\text{相異なる 5 個のものから 3 個を選んで並べたものの総数}}{\text{相異なる 3 つのものを並べたものの総数}}
$$

$$
= \frac{\text{5 から順に 1 ずつ減らしながら 3 つの整数をかけたもの}}{\text{3 から順に 1 ずつ減らしながら 3 つの整数をかけたもの}}
$$

$$
= \frac{5 \times 4 \times 3}{3 \times 2 \times 1} = 10
$$

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

以上より，「1」が出る回数が 10 である確率は，

　はじめの 10 回は「1」が出て残りの 50 回は「1 以外」が出る確率

$$\left(\frac{1}{6}\right)^{10} \times \left(\frac{5}{6}\right)^{50}$$

を ${}_{60}\mathrm{C}_{10}$ 個足し合わせることによって求められます
（加法定理（2 つの事象が互いに排反である場合））．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

つまり，

$$ {}_{60}\mathrm{C}_{10} \times \left(\frac{1}{6}\right)^{10} \times \left(\frac{5}{6}\right)^{50} $$

を計算することにより求められます．

Point! **サイコロを 60 回振ったときの 1 が出る回数が 10 である確率**

サイコロを 60 回振ったときの 1 の目が出る回数 X が 10 である確率 $P(X = 10)$ は，次で求められます．

$$P(X = 10) = {}_{60}\mathrm{C}_{10} \times \left(\frac{1}{6}\right)^{10} \times \left(\frac{5}{6}\right)^{50}$$

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

ここで，一般に，事象 A と事象 B が同時に起こることがない場合，互いに排反であるといわれ，次が成り立ちます．

加法定理（2 つの事象が互いに排反である場合）

事象 A と事象 B が互いに排反であるとき，次が成り立ちます．

$$A \text{ または } B \text{ が起こる確率} = A \text{ が起こる確率} + B \text{ が起こる確率}$$

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

F	G	H	I
理論→	「1」が出る確率		はじめの10回は「1」が出て残りの50回は「1以外」が出る確率
	0.166666667		1.81729E-12
	「1以外」が出る確率		
	0.833333333		「1」が出る回数が10であるときの「1」と「1以外」の並び方のパターンの総数
			75394027566
			「1」が出る回数が10である確率
			0.137013114

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

以上において,「サイコロを振ったときに 1 の目が出るか出ないか」という試行を考えました．これは，起こりうる結果が,「1 の目が出る」「1 以外の目が出る」の 2 つのみなので，ベルヌーイ試行です．

例題 4-1　サイコロを 60 回振ったときの 1 の目が出る割合を調べる

Point!

ベルヌーイ試行

「サイコロを振ったときに 1 の目が出るか出ないか」「コインを投げたときに表が出るか裏が出るか」のように，起こりうる結果が「成功」と「失敗」の 2 通りに限られる試行はベルヌーイ試行といわれます．

4.2　2項分布

例題 4-2　サイコロを 60 回振ったときの 1 の目が出る回数についての相対度数を調べる

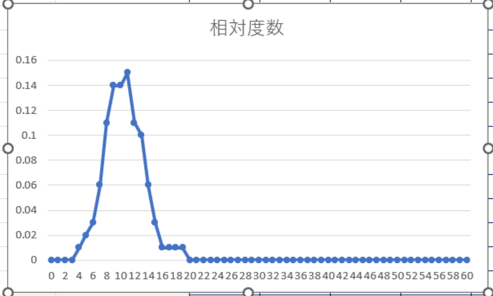
サイコロを 60 回振ったときの 1 の目が出る回数を X として，この試行を 100 回くりかえしたときの，$X = 0, 1, \cdots, 60$ についての度数および相対度数を計算してみましょう．

また，相対度数のマーカー付き折れ線グラフを作成しましょう．

例題 4-2　サイコロを 60 回振ったときの 1 の目が出る回数についての相対度数を調べる

	A	B	C	D	E
1					
2	「1」が出る回数X	度数	相対度数		確率
3	0	0	0		
4	1	0	0		
5	2	0	0		
6	3	0	0		
7	4	1	0.01		
8	5	2	0.02		
9	6	3	0.03		
10	7	6	0.06		
11	8	11	0.11		
12	9	14	0.14		
13	10	14	0.14		
14	11	15	0.15		
15	12	11	0.11		
16	13	10	0.1		
17	14	6	0.06		

	G	H	I	J	K	L	
1		試行1	試行2	試行3	試行4	試行5	試行
2	「1」が出る回数X	19	11	13	12	13	
4		6	5	4	3	6	
5						2	
6						2	
7						1	
8						1	
9						3	
10						5	
11						6	
12						5	
13						1	
14						1	
15						4	
16						1	
17		1	4	3	3	6	



Sheet1

例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

サイコロを 60 回振ったときの 1 の目が出る回数を X として，$X=1$ である確率 $P(X=1)$，$X=2$ である確率 $P(X=2)$，$\cdots$，$X=60$ である確率 $P(X=60)$ をそれぞれ計算してみましょう．

また,「相対度数のマーカー付き折れ線グラフ」と「確率分布の棒グラフ」の形がお互い似ていることを確かめましょう．

例題4-3　サイコロを60回振ったときの1の目が出る回数についての確率を求める

ここで，X は確率変数であり，$1,\ 2,\ \cdots,\ 60$ のどれかの値をとります．

例題4-1では，1の目が出る回数が10である確率を求め，約0.137であることがわかりました．これは，$P(X=10)$ が約0.137であると表現することができます．

例題4-3　サイコロを60回振ったときの1の目が出る回数についての確率を求める

Point! **確率変数**

たとえば，サイコロを振ったときに「1の目が出る」という事象に「1」を,「2の目が出る」という事象に「2」を，$\cdots$,「6の目が出る」という事象に「6」を割り当てるという対応を X とすると，X のとりうる値は $1,\ 2,\ 3,\ 4,\ 5,\ 6$ のどれかであり，それぞれの値の目が出る確率が決まっています.

例題4-3　サイコロを60回振ったときの1の目が出る回数についての確率を求める

Point! **確率変数**

1の目が出る確率が$1/6$であることを，$X = 1$である確率が$1/6$であるといい，

$$P(X = 1) = \frac{1}{6}$$

とあらわします．

例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

Point! 確率変数

この X は，各根元事象（「1 の目が出る」など）から数（1 など）への対応であり，X のとりうる値それぞれについて，それに対応する事象が起こる確率が決まっています．このような対応は確率変数とよばれます．

例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

確率変数は，これらの X のように，とりうる値がとびとびの値のときには離散型確率変数とよばれ，連続的な値（つながった値）のときには連続型確率変数とよばれます．

例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

Point!

離散型確率分布

離散型確率変数 X のとる値 x から，$X = x$ である確率 $P(X = x)$ への対応は，X の（離散型）確率分布といわれます．

「全事象が起こる確率は 1 である」ことから，それぞれの確率を全部足すと 1 になります．

例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

Point! 離散型確率分布

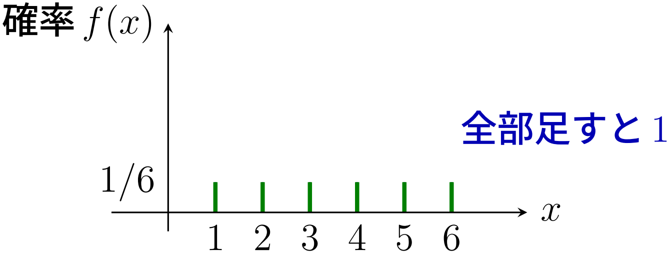
また，この対応を定める関数 f は，X の確率質量関数といわれます（$f(x) = P(X = x)$）．

x	1	2	3	4	5	6
$P(X = x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

Point!

離散型確率分布



離散型確率分布（サイコロを振ったときに出る目を X としている）

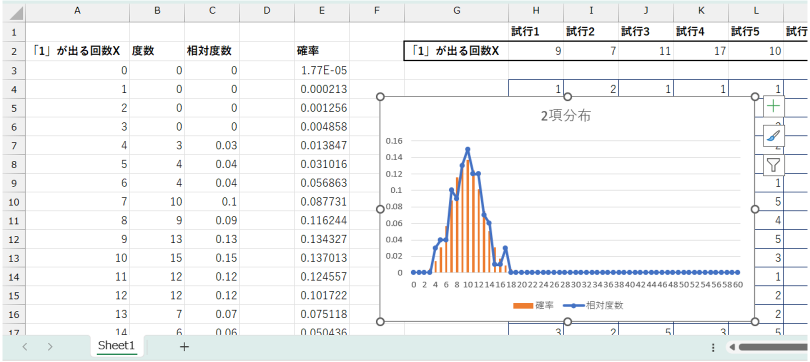
例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

Point! **サイコロを 60 回振ったときの 1 が出る回数が k である確率**

サイコロを 60 回振ったときの 1 の目が出る回数 X が k である確率 $P(X = k)$ は，次で求められます．

$$P(X = k) = {}_{60}\mathrm{C}_k \times \left(\frac{1}{6}\right)^k \times \left(\frac{5}{6}\right)^{(60-k)}$$

例題4-3　サイコロを60回振ったときの1の目が出る回数についての確率を求める



	A	B	C	D	E
1					
2	「1」が出る回数X	度数	相対度数		確率
3	0	0	0		1.77E-05
4	1	0	0		0.000213
5	2	0	0		0.001256
6	3	0	0		0.004858
7	4	3	0.03		0.013847
8	5	4	0.04		0.031016
9	6	4	0.04		0.056863
10	7	10	0.1		0.087731
11	8	9	0.09		0.116244
12	9	13	0.13		0.134327
13	10	15	0.15		0.137013
14	11	12	0.12		0.124557
15	12	12	0.12		0.101722
16	13	7	0.07		0.075118
17	14	6	0.06		0.050436

「相対度数のマーカー付き折れ線グラフ」と「確率分布の棒グラフ」の形がお互い似ていることが確かめられます.

例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

サイコロを 60 回振ったときの 1 の出る回数を X として，$X = 1$ である確率 $P(X = 1)$，$X = 2$ である確率 $P(X = 2)$，$\cdots$，$X = 60$ である確率 $P(X = 60)$ をそれぞれ計算しました．このような確率分布は **2 項分布**とよばれます．

また，この例の場合，X は 2 項分布 $B(60, 1/6)$ に従うといいます．
$B(60, 1/6)$ は試行回数 60，成功率 1/6 の 2 項分布をあらわします．

例題 4-3　サイコロを 60 回振ったときの 1 の目が出る回数についての確率を求める

2 項分布

ベルヌーイ試行を，成功率 p が一定で，各回における結果（成功または失敗）が互いに独立になるように n 回おこなったとき，成功数（成功する回数）$\boldsymbol{X}$ の確率分布は，試行回数 $\boldsymbol{n}$，成功率 $\boldsymbol{p}$ の 2 項分布とよばれ，$\boldsymbol{B(n, p)}$ と記されます．またこのとき，$\boldsymbol{X}$ は 2 項分布 $\boldsymbol{B(n, p)}$ に従うといわれます．

2 項分布は離散型確率分布です．

例題4-3　サイコロを60回振ったときの1の目が出る回数についての確率を求める

Point!

確率変数 X が2項分布に従うときの $X = k$ である確率

確率変数 X が試行回数 n，成功率 p の2項分布 $B(n, p)$ に従うとき，$X = k$ である確率 $P(X = k)$ は，次で求められます．

$$P(X = k) = {}_n\mathrm{C}_k \times p^k \times (1 - p)^{(n-k)}$$

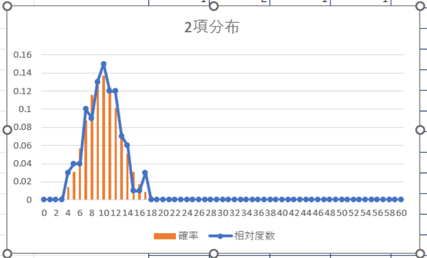
例題 4-4　2 項分布を Excel 関数で求める

例題 4-3 において，2 項分布 $B(60, 1/6)$ を Excel 関数で求めなおしてみましょう．そして，それが,（例題 4-3 で求めた）サイコロを 60 回振ったときの 1 の出る回数 X についての $X = 1$ である確率 $P(X = 1)$，$X = 2$ である確率 $P(X = 2)$，$\cdots$，$X = 60$ である確率 $P(X = 60)$ と一致することを確認しましょう．

例題 4-4　2 項分布を Excel 関数で求める

	A	B	C	D	E
1					
2	「1」が出る回数X	度数	相対度数		確率
3	0	0	0		1.77E-05
4	1	0	0		0.000213
5	2	0	0		0.001256
6	3	0	0		0.004858
7	4	3	0.03		0.013847
8	5	4	0.04		0.031016
9	6	4	0.04		0.056863
10	7	10	0.1		0.087731
11	8	9	0.09		0.116244
12	9	13	0.13		0.134327
13	10	15	0.15		0.137013
14	11	12	0.12		0.124557
15	12	12	0.12		0.101722
16	13	7	0.07		0.075118
17	14	6	0.06		0.050436

	G	H	I	J	K	L
1		試行1	試行2	試行3	試行4	試行5
2	「1」が出る回数X	9	7	11	17	10
4		1	2	1	1	1
9						1
10						5
11						4
12						5
13						3
14						1
15						2
16						2
17		3	2	5	3	5



Sheet1

BINOM.DIST 関数で求めなおしても，まったく値が変わらないことがわかります．

例題 4-5　確率分布の期待値を求める

例題 4-4 において，「**A** 列の値（「**1**」が出る回数 X）」と「**C** 列の値（相対度数）」の積和（つまり，「**1**」が出た回数 X の平均値），および，2 項分布 $B(60, 1/6)$ の期待値を求めてみましょう．

例題 4-5　確率分布の期待値を求める

ここで，「**A 列の値（「1」が出る回数 X）**」と「**C 列の値（相対度数）**」の積和とは

「A 列の値（「1」が出る回数 X）」と「C 列の値（相対度数）」の対応する値どうしの積すべての和

という意味であり，つまり，

「「1」が出る回数 X」と「度数/全データ数」の対応する値どうしの積すべての和

となります．

例題 4-5　確率分布の期待値を求める

なので，これは，

$$\frac{\text{「}1\text{ が出る回数 }X\text{」と「度数」の対応する値どうしの積すべての和}}{\text{全データ数}}$$

つまり，

「1」が出た回数 $\boldsymbol{X}$ の平均値

のことであることがわかります．

また，2 項分布 $B(60, 1/6)$ の期待値も同じように求められます．

例題 4-5　確率分布の期待値を求める

確率分布の期待値

確率変数 X の確率分布の期待値とは，X がとると「期待」される値のことです．また，これは X の期待値ともいわれます．

X が離散型確率変数のときは，「$\boldsymbol{X}$ がとる値」と「その値をとる確率」の積すべての和が，X の確率分布の期待値となります．

例題 4-5　確率分布の期待値を求める

なお，確率変数 X の期待値は，理論的な平均値，つまり,「X が従う確率分布をとるような母集団の平均値（母平均）」をあらわすもので，サイコロを「振る前」にも求められるようなものです．

一方，データの平均値（標本平均）は，サイコロを「振ったあと」に結果（標本）に対して求めるようなものです．

例題 4-5　確率分布の期待値を求める

60	[illegible]	[illegible]	[illegible]		[illegible]			[illegible]	[illegible]
61	58	0	0		9.05E-43			3	3
62	59	0	0		6.14E-45			5	6
63	60	0	0		2.05E-47			6	3
64									
65	**平均値（期待値）**		10.17		10				
66	**分散**								
67									

「1」が出た回数 X の平均値が計算されます．また，2 項分布 $B(60, 1/6)$ の期待値 10 が計算されます．

例題 4-6　確率分布の分散を求める

例題 4-5 において，「**A** 列の値（「**1**」が出る回数 X）と平均値の差の **2** 乗」と「**C** 列の値（相対度数）」の積和（つまり，「**1**」が出た回数 X の分散），および，2 項分布 $B(60, 1/6)$ の分散を求めてみましょう．

例題 4-6　確率分布の分散を求める

ここで，**「A 列の値（「1」が出る回数 X）と平均値の差の 2 乗」と「C 列の値（相対度数）」の積和とは**

「A 列の値（「1」が出る回数 X）と平均値の差の 2 乗」と「C 列の値（相対度数）」の対応する値どうしの積すべての和

という意味であり，つまり，

「「1」が出る回数 X と平均値の差の 2 乗」と「度数/全データ数」の対応する値どうしの積すべての和

となります．

例題 4-6　確率分布の分散を求める

なので，これは，

$$\frac{\text{「1 が出る回数}\ X\ \text{と平均値の差の 2 乗」と「度数」の対応する値どうしの積すべての和}}{\text{全データ数}}$$

つまり，

「1」が出た回数 X の分散

のことであることがわかります．

また，2 項分布 $B(60, 1/6)$ の分散も同じように求められます．

例題 4-6　確率分布の分散を求める

Point! 確率分布の分散，標準偏差

確率変数 X について,「X がとる値と期待値の差の 2 乗」の期待値は，X の確率分布の分散とよばれます．また，これは X の分散ともいわれます．

分散の正の平方根は標準偏差とよばれます．

X が離散型確率変数のときは，「X がとる値と期待値の差の 2 乗」と「その値をとる確率」の積すべての和が，X の確率分布の分散となります．

例題 4-6　確率分布の分散を求める

確率変数の分散は，確率変数の値が期待値を中心としてどれくらいばらついて分布しているかをあらわす指標です．つまり,「分布のばらつきの大きささをあらわす数値」だといえます．

例題 4-6　確率分布の分散を求める

なお，確率変数 X の分散は，理論的な分散，つまり,「X が従う確率分布をとるような母集団の分散（母分散）」をあらわすもので，サイコロを「振る前」にも求められるようなものです．

一方，データの分散（標本分散）は，サイコロを「振ったあと」に結果（標本）に対して求めるようなものです．

例題 4-6　確率分布の分散を求める

60	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		[illegible]	[illegible]
61	58	0	0	2336.756	9.05E-43	2304		5	4
62	59	0	0	2434.436	6.14E-45	2401		6	3
63	60	0	0	2534.116	2.05E-47	2500		1	6
64									
65	平均値（期待値）		9.66		10				
66	分散		7.7644		8.333333				
67									

「1」が出た回数 X の分散が計算されます．また，2 項分布 $B(60, 1/6)$ の分散 $25/3$ $(= 8.33\cdots)$ が計算されます．

例題 4-7　2 項分布の期待値と分散を求める

例題 4-6 において，試行回数 60，成功率 1/6 の 2 項分布 $B(60, 1/6)$ の期待値を

「試行回数 × 成功率」

を計算することにより求めなおしましょう．また，分散を

「試行回数 × 成功率 × 失敗率」

つまり「試行回数 × 成功率 ×(1− 成功率)」

を計算することにより求めなおしましょう．

例題 4-7　2 項分布の期待値と分散を求める

ここで，一般に次が成り立ちます．

Point! 2 項分布の期待値と分散

確率変数 X が試行回数 n，成功率 p の 2 項分布 $B(n,p)$ に従うとき，X の期待値 $E(X)$ と分散 $V(X)$ は次のように求められます．

$$E(X) = np, \quad V(X) = np(1-p)$$

例題 4-7　2 項分布の期待値と分散を求める

「=60*(1/6)」と入力し，$B(60, 1/6)$ の期待値を計算すると，例題 4-5 で求めた結果 10 と一致することがわかります．

また,「=60*(1/6)*(1-1/6)」と入力し，$B(60, 1/6)$ の分散を計算すると，例題 4-6 で求めた結果 $8.33\cdots$ と一致することがわかります．

第 5 章　正規分布

5.1　2 項分布から正規分布へ

例題 5-1　2 項分布において試行回数を増やす

サイコロを 600 回振ったときの 1 の目が出る回数 X の確率分布，つまり，2 項分布 $B(600, 1/6)$ の棒グラフを作成しましょう．

例題 5-1　2 項分布において試行回数を増やす

ここで，2 項分布は，たとえばサイコロを振る試行やコインを投げる試行のような具体的な状況を確率の観点からモデル化したもので，試行回数 n と成功率 p によって分布が決まります．

前章では，試行回数 60，成功率 $1/6$ の 2 項分布 $B(60, 1/6)$ の棒グラフを作成しました．横軸の確率変数は離散型です．つまり，とびとびの値をとり，連続的な値（つながった値）をとりません．

例題 5-1　2 項分布において試行回数を増やす

	A	B	C	D	E
1					
2	「1」が出る回数X	度数	相対度数		確率
3	0	0	0		1.77E-05
4	1	0	0		0.000213
5	2	0	0		0.001256
6	3	0	0		0.004858
7	4	3	0.03		0.013847
8	5	4	0.04		0.031016
9	6	4	0.04		0.056863
10	7	10	0.1		0.087731
11	8	9	0.09		0.116244
12	9	13	0.13		0.134327
13	10	15	0.15		0.137013
14	11	12	0.12		0.124557
15	12	12	0.12		0.101722
16	13	7	0.07		0.075118
17	14	6	0.06		0.050436

	F	G	H	I	J	K	L
1			試行1	試行2	試行3	試行4	試行5
2		「1」が出る回数X	9	7	11	17	10
4			1	2	1	1	1
9							1
10							5
11							4
12							5
13							3
14							1
15							2
16							2
17			3	2	5	3	5

Sheet1

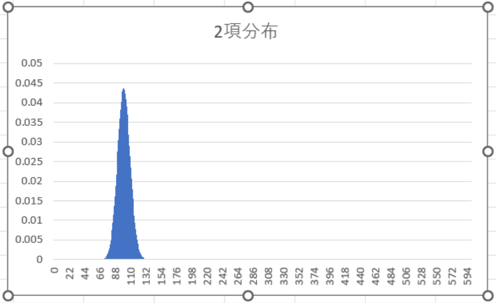
2 項分布 $B(60, 1/6)$ の棒グラフ（例題 4-3）

例題 5-1　2 項分布において試行回数を増やす

ここでは，試行回数 n を 60 から 600 へと大きくしてみると，2 項分布はどうなるのか確認しましょう．

例題 5-1　2 項分布において試行回数を増やす

	A	B
1	「1」が出る回数X	2項分布
2	0	3.1E-48
3	1	3.72E-46
4	2	2.23E-44
5	3	8.88E-43
6	4	2.65E-41
7	5	6.32E-40
8	6	1.25E-38
9	7	2.13E-37
10	8	3.15E-36
11	9	4.15E-35
12	10	4.9E-34
13	11	5.26E-33
14	12	5.16E-32
15	13	4.67E-31
16	14	3.92E-30
17	15	3.06E-29



Sheet1

例題 5-1　2 項分布において試行回数を増やす

試行回数が 60 では，ばらばらの棒グラフの集まりにみえた 2 項分布が，試行回数が増えて 600 になると，棒グラフがつながってみえます．

このように，2 項分布において試行回数 n を大きくすると，なめらかな連続するグラフに近づいていきます．この連続するグラフは正規分布のグラフであり，これを与える関数は，正規分布の確率密度関数とよばれます．

例題 5-1　2 項分布において試行回数を増やす

確率変数が離散的な値（とびとびの値）をとる 2 項分布から，確率変数が連続的な値（つながった値）をとる確率分布への一般化として，正規分布が定式化されます．

例題 5-1　2 項分布において試行回数を増やす

Point!

連続型確率分布

連続型確率変数 X においては，たとえば，

$$P(0.2 \leq X \leq 0.5) = 0.3$$

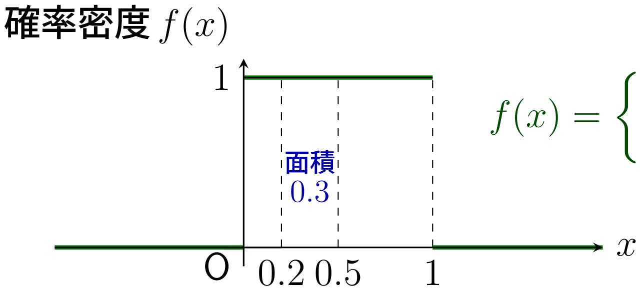
（0.2 から 0.5 の区間に入る確率は 0.3）

のように，区間について確率を考えます．

例題 5-1　2 項分布において試行回数を増やす

Point!

連続型確率分布



$$f(x) = \begin{cases} 1 & (0 \leq x \leq 1) \\ 0 & (\text{それ以外}) \end{cases}$$

連続型確率分布（区間 $[0.2, 0.5]$ に入る確率は 0.3，区間 $(-\infty, \infty)$ に入る確率は 1）

例題 5-1　2 項分布において試行回数を増やす

（連続型）確率分布

このような確率を，上図の f のように，その範囲についてのグラフの下側の面積を計算することにより求めることができるように定められた関数は，確率密度関数といわれます．

例題 5-1　2 項分布において試行回数を増やす

Point! 連続型確率分布

連続型確率変数 X において，その確率密度関数によって与えられる対応は（連続型）確率分布といわれます．

例題 5-1　2 項分布において試行回数を増やす

Point! **連続型確率分布**

連続型確率分布のグラフにおいて，横軸は確率変数 X がとる値をあらわしますが，縦軸があらわしているものは X がとる値の「相対的な起こりやすさ」であり，確率密度といわれます．

例題 5-1　2 項分布において試行回数を増やす

Point!

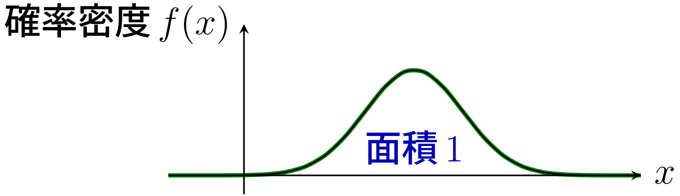
連続型確率分布

連続型確率変数 X については，X がある値 x をとるときの確率密度は確率をあらわしませんが，ある区間を考えれば，グラフの下側の面積が確率をあらわします．

「全事象が起こる確率は 1 である」ことから，全区間についてのグラフの下側の面積は 1 になります．

例題 5-1　2 項分布において試行回数を増やす

Point! **連続型確率分布**



連続型確率分布（区間 $(-\infty, \infty)$ に入る確率は 1）

正規分布は連続型確率分布です．

例題 5-1　2 項分布において試行回数を増やす

Point! **正規分布**

期待値 μ，分散 σ^2（標準偏差 σ）とする正規分布は，確率密度関数が

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

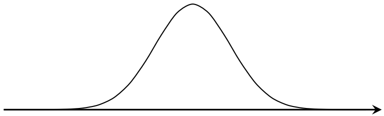
の形で与えられる連続型確率分布であり，μ と σ^2 を決めればただ 1 つに定まります．

この確率分布を $\boldsymbol{N(\mu, \sigma^2)}$ と書きます．そして，確率変数 X の確率分布が正規分布 $N(\mu, \sigma^2)$ であるとき，X は正規分布 $\boldsymbol{N(\mu, \sigma^2)}$ に従うといいます．

例題 5-1　2 項分布において試行回数を増やす

Point!

正規分布



試行回数 n，成功率 p の 2 項分布 $B(n, p)$ は，試行回数 n が大きくなるにつれて，期待値 np，分散 $np(1-p)$ の正規分布 $N(np, np(1-p))$ に近づきます．

例題 5-1　2 項分布において試行回数を増やす

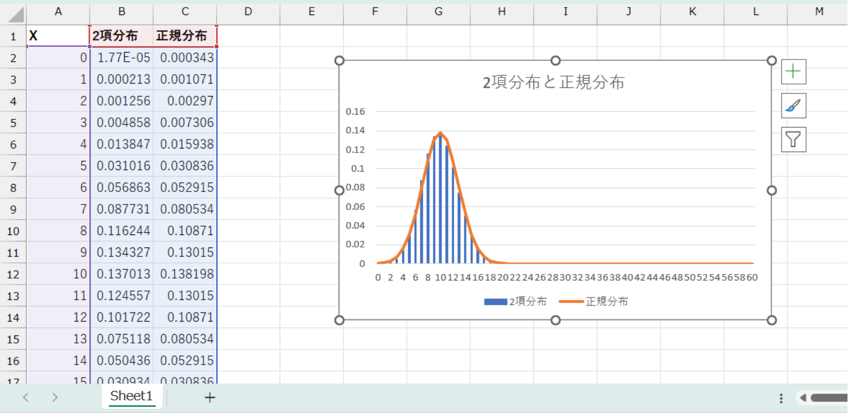
正規分布は期待値を中心とする左右対称な連続型確率分布です．

統計学における推定や検定，自然科学や社会科学におけるモデルの作成など，さまざまな場面で活用されています．

例題 5-2　2 項分布の近似曲線を作成する

サイコロを 60 回振ったときの 1 の目が出る回数 X についての確率分布，つまり，2 項分布 $B(60, 1/6)$ のグラフが正規分布 $N(10, 25/3)$ のグラフで近似されていることを確かめましょう．

例題 5-2　2 項分布の近似曲線を作成する



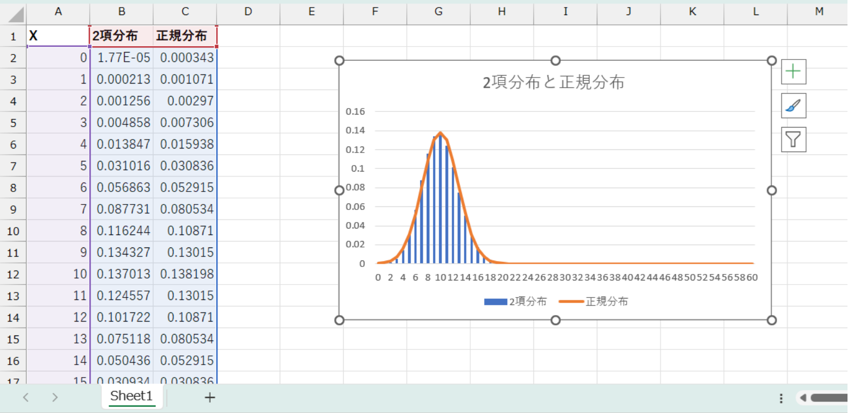
X	2項分布	正規分布
0	1.77E-05	0.000343
1	0.000213	0.001071
2	0.001256	0.00297
3	0.004858	0.007306
4	0.013847	0.015938
5	0.031016	0.030836
6	0.056863	0.052915
7	0.087731	0.080534
8	0.116244	0.10871
9	0.134327	0.13015
10	0.137013	0.138198
11	0.124557	0.13015
12	0.101722	0.10871
13	0.075118	0.080534
14	0.050436	0.052915
15	0.030934	0.030836

2 項分布 $B(60, 1/6)$ のグラフが正規分布 $N(10, 25/3)$ のグラフで近似されていることが確かめられます．

例題 5-3　正規分布を Excel 関数で求める

例題 5-2 において，2 項分布 $B(60, 1/6)$ のグラフに近似する正規分布 $N(10, 25/3)$ を Excel 関数で求めてみましょう．そして，それが, (例題 5-2 で求めた）正規分布の確率密度関数と一致することを確認しましょう.

例題 5-3　正規分布を Excel 関数で求める



X	2項分布	正規分布
0	1.77E-05	0.000343
1	0.000213	0.001071
2	0.001256	0.00297
3	0.004858	0.007306
4	0.013847	0.015938
5	0.031016	0.030836
6	0.056863	0.052915
7	0.087731	0.080534
8	0.116244	0.10871
9	0.134327	0.13015
10	0.137013	0.138198
11	0.124557	0.13015
12	0.101722	0.10871
13	0.075118	0.080534
14	0.050436	0.052915
15	0.030934	0.030836

NORM.DIST 関数で求めなおしても，まったく値が変わらないことがわかります．

5.2　累積分布関数

例題 5-4　正規分布の累積分布関数を求める

ある試験の点数 X が期待値 50，標準偏差 10 の正規分布 $N(50, 10^2)$ に従うとして，X の確率分布のグラフを作成しましょう．

また，X の累積分布関数のグラフも作成しましょう．

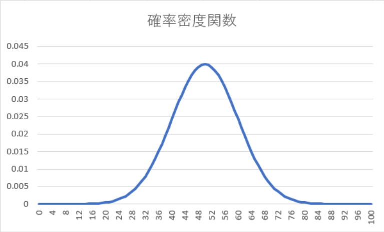
例題 5-4　正規分布の累積分布関数を求める

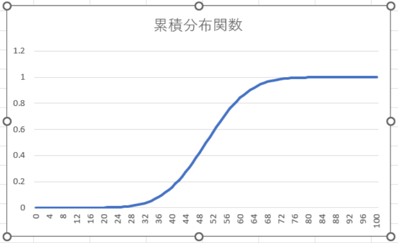
Point! **累積分布関数**

確率変数 X がある値 x 以下になる確率 $P(X \leq x)$ をあらわす関数は，累積分布関数といわれます．

例題 5-4　正規分布の累積分布関数を求める

	A	B	C
1	点数X	確率密度関数	累積分布関数
2	0	1.48672E-07	2.86652E-07
3	1	2.43896E-07	4.79183E-07
4	2	3.9613E-07	7.93328E-07
5	3	6.36983E-07	1.30081E-06
6	4	1.01409E-06	2.11245E-06
7	5	1.59837E-06	3.39767E-06
8	6	2.49425E-06	5.41254E-06
9	7	3.85352E-06	8.53991E-06
10	8	5.89431E-06	1.33457E-05
11	9	8.92617E-06	2.06575E-05
12	10	1.3383E-05	3.16712E-05
13	11	1.98655E-05	4.80963E-05
14	12	2.91947E-05	7.2348E-05
15	13	4.2478E-05	0.0001078
16	14	6.11902E-05	0.000159109
17	15	8.72683E-05	0.000232629





	E	H
14	30≦X≦70	2σ区間
15	20≦X≦80	3σ区間
16	10≦X≦90	4σ区間

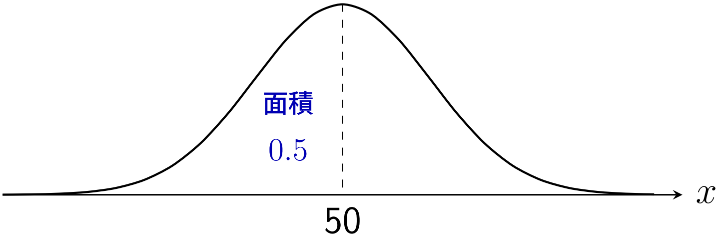
Sheet1

例題 5-5　正規分布についてある値以下になる確率，ある値以上になる確率を求める

ある試験の点数 X が期待値 50，標準偏差 10 の正規分布 $N(50, 10^2)$ に従うとして，次をそれぞれ求めましょう．

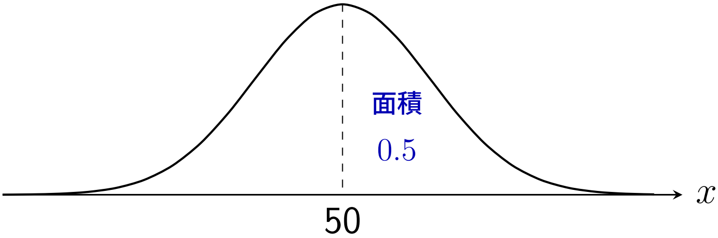
$$P(X \leq 50)，P(X \geq 50)，P(X \leq 40)，P(X \geq 60)，$$
$$P(X \leq 30)，P(X \geq 70)，P(X \leq 20)，P(X \geq 80)，$$
$$P(X \leq 10)，P(X \geq 90)$$

例題 5-5　正規分布についてある値以下になる確率，ある値以上になる確率を求める



正規分布 $N(50, 10^2)$ の確率密度関数のグラフ（$X \leq 50$ である確率は 0.5）

例題 5-5　正規分布についてある値以下になる確率，ある値以上になる確率を求める

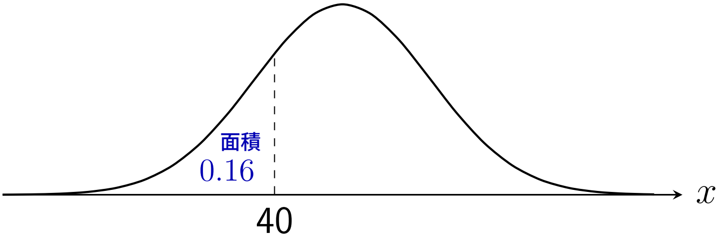


正規分布 $N(50, 10^2)$ の確率密度関数のグラフ（$X \geq 50$ である確率は 0.5）

例題 5-5　正規分布についてある値以下になる確率，ある値以上になる確率を求める

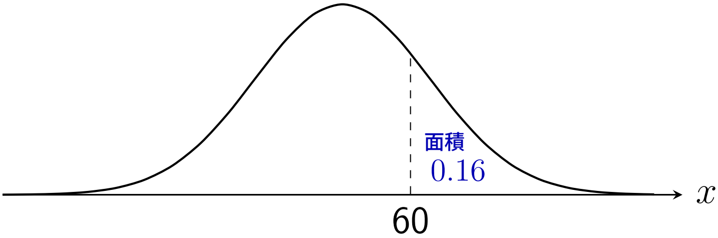
一般に，正規分布は期待値を中心として左右対称なので，それに従う確率変数 X について，X の値が期待値以下となる確率も，期待値以上になる確率も，0.5 になります．

例題 5-5　正規分布についてある値以下になる確率，ある値以上になる確率を求める



正規分布 $N(50, 10^2)$ の確率密度関数のグラフ（$X \leq 40$ である確率は約 0.16）

例題 5-5　正規分布についてある値以下になる確率，ある値以上になる確率を求める



正規分布 $N(50, 10^2)$ の確率密度関数のグラフ（$X \geq 60$ である確率は約 0.16）

例題 5-5　正規分布についてある値以下になる確率，ある値以上になる確率を求める

	A	B	C	D	E	F	G	H
1	点数X	確率密度関数	累積分布関数					
2	0	1.48672E-07	2.86652E-07		区間	確率	人数	
3	1	2.43896E-07	4.79183E-07		X≦50	0.5		
4	2	3.9613E-07	7.93328E-07		X≧50	0.5		
5	3	6.36983E-07	1.30081E-06		X≦40	0.158655		
6	4	1.01409E-06	2.11245E-06		X≧60	0.158655		
7	5	1.59837E-06	3.39767E-06		X≦30	0.02275		
8	6	2.49425E-06	5.41254E-06		X≧70	0.02275		
9	7	3.85352E-06	8.53991E-06		X≦20	0.00135		
10	8	5.89431E-06	1.33457E-05		X≧80	0.00135		
11	9	8.92617E-06	2.06575E-05		X≦10	3.17E-05		
12	10	1.3383E-05	3.16712E-05		X≧90	3.17E-05		
13	11	1.98655E-05	4.80963E-05		40≦X≦60			1σ区間
14	12	2.91947E-05	7.2348E-05		30≦X≦70			2σ区間
15	13	4.2478E-05	0.0001078		20≦X≦80			3σ区間
16	14	6.11902E-05	0.000159109		10≦X≦90			4σ区間
17	15	8.72683E-05	0.000232629					

Sheet1

例題 5-6　正規分布について 1σ 区間，2σ 区間，3σ 区間に入る確率を求める

ある試験の点数 X が期待値 50，標準偏差 10 の正規分布 $N(50, 10^2)$ に従うとして，次をそれぞれ求めましょう．

$P(40 \le X \le 60)$，$P(30 \le X \le 70)$，

$P(20 \le X \le 80)$，$P(10 \le X \le 90)$

例題 5-6　正規分布について σ 区間に入る確率を求める

Point!

$\boldsymbol{1\sigma}$ 区間，$\boldsymbol{2\sigma}$ 区間，$\boldsymbol{3\sigma}$ 区間，$\boldsymbol{4\sigma}$ 区間

標準偏差 σ の正規分布において，期待値から（プラスマイナス）σ の範囲，2σ の範囲，3σ の範囲，4σ の範囲をそれぞれ，

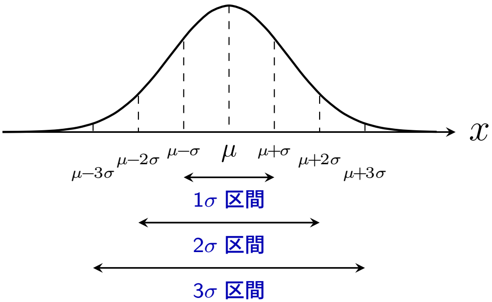
$\boldsymbol{1\sigma}$ 区間，$\boldsymbol{2\sigma}$ 区間，$\boldsymbol{3\sigma}$ 区間，$\boldsymbol{4\sigma}$ 区間

といいます．

例題 5-6　正規分布について σ 区間に入る確率を求める

Point!

1σ 区間，2σ 区間，3σ 区間，4σ 区間



正規分布 $N(\mu, \sigma^2)$ の確率密度関数のグラフ（1σ 区間，2σ 区間，3σ 区間）

例題 5-6　正規分布について σ 区間に入る確率を求める

Point! **1σ 区間，2σ 区間，3σ 区間，4σ 区間**

たとえば，X が期待値 50，標準偏差 10 の正規分布に従うときは，次のようになります．

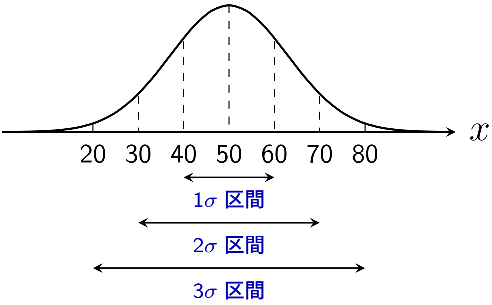
1σ 区間は $40 \leq X \leq 60$，　2σ 区間は $30 \leq X \leq 70$，

3σ 区間は $20 \leq X \leq 80$，　4σ 区間は $10 \leq X \leq 90$

例題 5-6　正規分布について σ 区間に入る確率を求める

Point!

1σ 区間，2σ 区間，3σ 区間，4σ 区間



正規分布 $N(50, 10^2)$ の確率密度関数のグラフ（1σ 区間，2σ 区間，3σ 区間）

例題 5-6　正規分布について σ 区間に入る確率を求める

	A	B	C	D	E	F	G	H
1	点数X	確率密度関数	累積分布関数					
2	0	1.48672E-07	2.86652E-07		区間	確率	人数	
3	1	2.43896E-07	4.79183E-07		X≦50	0.5		
4	2	3.9613E-07	7.93328E-07		X≧50	0.5		
5	3	6.36983E-07	1.30081E-06		X≦40	0.158655		
6	4	1.01409E-06	2.11245E-06		X≧60	0.158655		
7	5	1.59837E-06	3.39767E-06		X≦30	0.02275		
8	6	2.49425E-06	5.41254E-06		X≧70	0.02275		
9	7	3.85352E-06	8.53991E-06		X≦20	0.00135		
10	8	5.89431E-06	1.33457E-05		X≧80	0.00135		
11	9	8.92617E-06	2.06575E-05		X≦10	3.17E-05		
12	10	1.3383E-05	3.16712E-05		X≧90	3.17E-05		
13	11	1.98655E-05	4.80963E-05		40≦X≦60	0.682689		1σ区間
14	12	2.91947E-05	7.2348E-05		30≦X≦70	0.9545		2σ区間
15	13	4.2478E-05	0.0001078		20≦X≦80	0.9973		3σ区間
16	14	6.11902E-05	0.000159109		10≦X≦90	0.999937		4σ区間
17	15	8.72683E-05	0.000232629					

Sheet1

例題 5-6　正規分布について σ 区間に入る確率を求める

Point! **1σ 区間，2σ 区間，3σ 区間，4σ 区間に入る確率**

標準偏差 σ の正規分布において，1σ 区間，2σ 区間，3σ 区間，4σ 区間に入る確率は，次のようになります．

1σ 区間に入る確率は約 68.27%，　2σ 区間に入る確率は約 95.45%，

3σ 区間に入る確率は約 99.73%，　4σ 区間に入る確率は約 99.99%

例題 5-7　正規分布について σ 区間に入るデータの個数を求める

ある試験の点数 X が期待値 50，標準偏差 10 の正規分布 $N(50, 10^2)$ に従うとして，次をそれぞれ求めましょう．ここで，この試験を受けた人数は 1000 とします．

例題 5-7　正規分布について σ 区間に入るデータの個数を求める

点数が 50 以下となる人数，点数が 50 以上となる人数，
点数が 40 以下となる人数，点数が 60 以上となる人数，
点数が 30 以下となる人数，点数が 70 以上となる人数，
点数が 20 以下となる人数，点数が 80 以上となる人数，
点数が 10 以下となる人数，点数が 90 以上となる人数，
点数が 40 以上 60 以下となる人数，点数が 30 以上 70 以下となる人数，
点数が 20 以上 80 以下となる人数，点数が 10 以上 90 以下となる人数

例題 5-7　正規分布について σ 区間に入るデータの個数を求める

たとえば，点数が 50 以下となる人数は，

「点数が **50** 以下となる確率」$\times$「全人数」

を計算して求めます．

例題 5-7　正規分布について σ 区間に入るデータの個数を求める

	A	B	C	D	E	F	G	H
1	点数X	確率密度関数	累積分布関数					
2	0	1.48672E-07	2.86652E-07		区間	確率	人数	
3	1	2.43896E-07	4.79183E-07		X≦50	0.5	500	
4	2	3.9613E-07	7.93328E-07		X≧50	0.5	500	
5	3	6.36983E-07	1.30081E-06		X≦40	0.158655	158.6553	
6	4	1.01409E-06	2.11245E-06		X≧60	0.158655	158.6553	
7	5	1.59837E-06	3.39767E-06		X≦30	0.02275	22.75013	
8	6	2.49425E-06	5.41254E-06		X≧70	0.02275	22.75013	
9	7	3.85352E-06	8.53991E-06		X≦20	0.00135	1.349898	
10	8	5.89431E-06	1.33457E-05		X≧80	0.00135	1.349898	
11	9	8.92617E-06	2.06575E-05		X≦10	3.17E-05	0.031671	
12	10	1.3383E-05	3.16712E-05		X≧90	3.17E-05	0.031671	
13	11	1.98655E-05	4.80963E-05		40≦X≦60	0.682689	682.6895	1σ区間
14	12	2.91947E-05	7.2348E-05		30≦X≦70	0.9545	954.4997	2σ区間
15	13	4.2478E-05	0.0001078		20≦X≦80	0.9973	997.3002	3σ区間
16	14	6.11902E-05	0.000159109		10≦X≦90	0.999937	999.9367	4σ区間
17	15	8.72683E-05	0.000232629					

Sheet1

第 6 章　標準正規分布

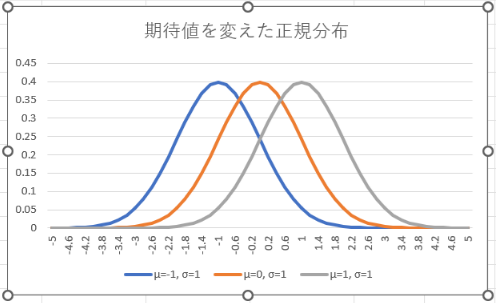
6.1　正規分布のグラフ

例題6-1　期待値を変えた正規分布のグラフを作成する

正規分布 $N(-1,1^2)$ のグラフ，$N(0,1^2)$ のグラフ，および，$N(1,1^2)$ のグラフを，同一グラフエリア内にそれぞれ作成して比べてみましょう．

例題 6-1　期待値を変えた正規分布のグラフを作成する

	A	B	C	D
1	X	μ=-1, σ=1	μ=0, σ=1	μ=1, σ=1
2	-5	0.00013383	1.48672E-06	6.07588E-09
3	-4.8	0.000291947	3.9613E-06	1.97732E-08
4	-4.6	0.000611902	1.01409E-05	6.18262E-08
5	-4.4	0.001232219	2.49425E-05	1.85736E-07
6	-4.2	0.002384088	5.89431E-05	5.36104E-07
7	-4	0.004431848	0.00013383	1.48672E-06
8	-3.8	0.007915452	0.000291947	3.9613E-06
9	-3.6	0.013582969	0.000611902	1.01409E-05
10	-3.4	0.02239453	0.001232219	2.49425E-05
11	-3.2	0.035474593	0.002384088	5.89431E-05
12	-3	0.053990967	0.004431848	0.00013383
13	-2.8	0.078950158	0.007915452	0.000291947
14	-2.6	0.110920835	0.013582969	0.000611902
15	-2.4	0.149727466	0.02239453	0.001232219
16	-2.2	0.194186055	0.035474593	0.002384088
17	-2	0.241970725	0.053990967	0.004431848



Sheet1

例題 6-1　期待値を変えた正規分布のグラフを作成する

正規分布 $N(-1, 1^2)$ のグラフを，右に（x 方向に）1 だけずらすと，正規分布 $N(0, 1^2)$ のグラフになり，これをさらに，右に（x 方向に）1 だけずらすと，正規分布 $N(1, 1^2)$ のグラフになります．

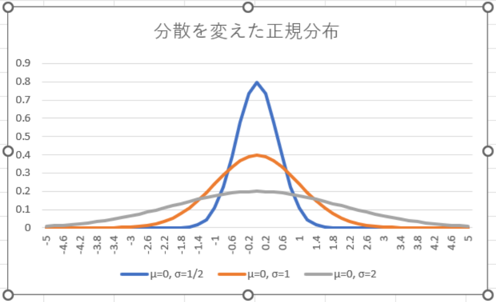
期待値を変えても分散はそのままなので，山の形は変わりません．そして，期待値が 1 大きくなると，山が 1 だけ右に移動することが確認できます．

例題6-2　分散を変えた正規分布のグラフを作成する

正規分布 $N(0,(1/2)^2)$ のグラフ，$N(0,1^2)$ のグラフ，および，$N(0,2^2)$ のグラフを，同一グラフエリア内にそれぞれ作成して比べてみましょう．

例題6-2　分散を変えた正規分布のグラフを作成する

	A	B	C	D
1	X	μ=0, σ=1/2	μ=0, σ=1	μ=0, σ=2
2	-5	1.53892E-22	1.48672E-06	0.00876415
3	-4.8	7.75622E-21	3.9613E-06	0.011197265
4	-4.6	3.33118E-19	1.01409E-05	0.014163519
5	-4.4	1.21915E-17	2.49425E-05	0.017737296
6	-4.2	3.80216E-16	5.89431E-05	0.021991798
7	-4	1.01045E-14	0.00013383	0.026995483
8	-3.8	2.28831E-13	0.000291947	0.032807907
9	-3.6	4.41598E-12	0.000611902	0.039475079
10	-3.4	7.26192E-11	0.001232219	0.047024539
11	-3.2	1.01763E-09	0.002384088	0.055460417
12	-3	1.21518E-08	0.004431848	0.064758798
13	-2.8	1.23652E-07	0.007915452	0.074863733
14	-2.6	1.07221E-06	0.013582969	0.085684296
15	-2.4	7.9226E-06	0.02239453	0.097093027
16	-2.2	4.98849E-05	0.035474593	0.108926089
17	-2	0.00026766	0.053990967	0.120985362



Sheet1

例題6-2　分散を変えた正規分布のグラフを作成する

分散を変えても期待値はそのままなので，山の位置は変わりません．そして，分散が小さくなると山が高く，すそがうすくなり，分散が大きくなると山が低く，すそが厚くなることが確認できます．

6.2　標準正規分布

例題 6-3　標準正規分布の累積分布関数を求める

確率変数 Z が標準正規分布 $N(0,1^2)$ に従うとして，Z の確率密度関数，累積分布関数の値をそれぞれ求めましょう．また，次をそれぞれ求めましょう．

$P(Z \leq 0)$，$P(Z \geq 0)$，$P(Z \leq -1)$，$P(Z \geq 1)$，

$P(Z \leq -2)$，$P(Z \geq 2)$，$P(Z \leq -3)$，$P(Z \geq 3)$，

$P(Z \leq -4)$，$P(Z \geq 4)$，

$P(-1 \leq Z \leq 1)$，$P(-1 \leq Z \leq 1)$，$P(-3 \leq Z \leq 3)$，$P(-4 \leq Z \leq 4)$

例題 6-3　標準正規分布の累積分布関数を求める

Point!

標準正規分布

期待値 0，分散 1（標準偏差 1）の正規分布 $N(0, 1^2)$ のことを標準正規分布といいます．

例題 6-3　標準正規分布の累積分布関数を求める

	A	B	C	D	E	F	G
1	Z	確率密度関数	累積分布関数		標準正規分布		
2	-5	1.48672E-06	2.86652E-07		区間	確率	
3	-4.9	2.43896E-06	4.79183E-07		Z≦0	0.5	
4	-4.8	3.9613E-06	7.93328E-07		Z≧0	0.5	
5	-4.7	6.36983E-06	1.30081E-06		Z≦-1	0.158655	
6	-4.6	1.01409E-05	2.11245E-06		Z≧1	0.158655	
7	-4.5	1.59837E-05	3.39767E-06		Z≦-2	0.02275	
8	-4.4	2.49425E-05	5.41254E-06		Z≧2	0.02275	
9	-4.3	3.85352E-05	8.53991E-06		Z≦-3	0.00135	
10	-4.2	5.89431E-05	1.33457E-05		Z≧3	0.00135	
11	-4.1	8.92617E-05	2.06575E-05		Z≦-4	3.17E-05	
12	-4	0.00013383	3.16712E-05		Z≧4	3.17E-05	
13	-3.9	0.000198655	4.80963E-05		-1≦Z≦1	0.682689	1σ区間
14	-3.8	0.000291947	7.2348E-05		-2≦Z≦2	0.9545	2σ区間
15	-3.7	0.00042478	0.0001078		-3≦Z≦3	0.9973	3σ区間
16	-3.6	0.000611902	0.000159109		-4≦Z≦4	0.999937	4σ区間
17	-3.5	0.000872683	0.000232629				

Sheet1

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

確率変数 Z が標準正規分布 $N(0, 1^2)$ に従うとして，次をそれぞれ求めましょう．

下位 50%以下に入るのはどの値以下か，

上位 50%以上に入るのはどの値以上か，

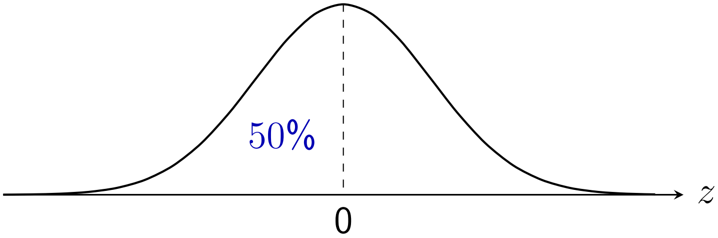
下位 15.86553%以下に入るのはどの値以下か，

上位 15.86553%以上に入るのはどの値以上か，

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

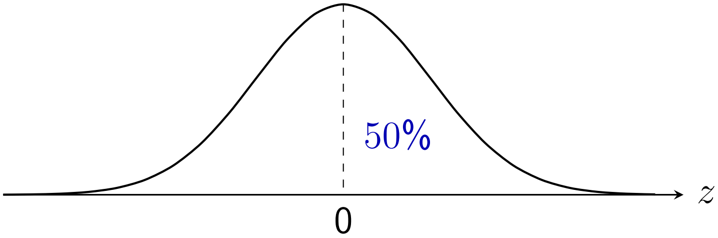
下位 5%以下に入るのはどの値以下か，
上位 5%以上に入るのはどの値以上か，
下位 2.5%以下に入るのはどの値以下か，
上位 2.5%以上に入るのはどの値以上か，
下位 0.5%以下に入るのはどの値以下か，
上位 0.5%以上に入るのはどの値以上か

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める



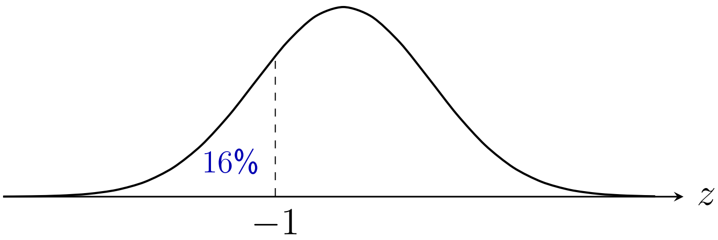
標準正規分布 $N(0,1^2)$ の確率密度関数のグラフ（下位 50%以下に入るのは 0 以下）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める



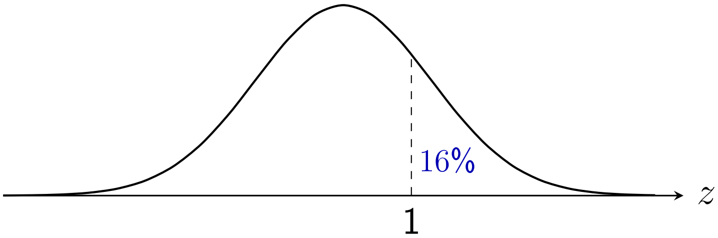
標準正規分布 $N(0,1^2)$ の確率密度関数のグラフ（上位 50%以上に入るのは 0 以上）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める



標準正規分布 $N(0,1^2)$ の確率密度関数のグラフ（下位 15.86553%以下に入るのは -1 以下）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（上位 15.86553%以上に入るのは 1 以上）

例題6-4　標準正規分布の累積分布関数の逆関数の値を求める

	A	B	C	D	E	F	G	H	I	J	K	L
1	Z	確率密度関数	累積分布関数		標準正規分布							
2	-5	1.48672E-06	2.86652E-07		下位50%以下に入るのはどの値以下か				0		P(Z≦0)	0.5
3	-4.8	3.9613E-06	7.93328E-07		上位50%以上に入るのはどの値以上か				0		P(Z≧0)	0.5
4	-4.6	1.01409E-05	2.11245E-06		下位15.86553%以下に入るのはどの値以下か				-1		P(Z≦-1)	0.1586553
5	-4.4	2.49425E-05	5.41254E-06		上位15.86553%以上に入るのはどの値以上か				1		P(Z≧1)	0.1586553
6	-4.2	5.89431E-05	1.33457E-05		下位5%以下に入るのはどの値以下か				-1.64485			
7	-4	0.00013383	3.16712E-05		上位5%以上に入るのはどの値以上か				1.644854			
8	-3.8	0.000291947	7.2348E-05		下位2.5%以下に入るのはどの値以下か				-1.95996			
9	-3.6	0.000611902	0.000159109		上位2.5%以上に入るのはどの値以上か				1.959964			
10	-3.4	0.001232219	0.000336929		下位0.5%以下に入るのはどの値以下か				-2.57583			
11	-3.2	0.002384088	0.000687138		上位0.5%以上に入るのはどの値以上か				2.575829			
12	-3	0.004431848	0.001349898		下位5%と上位5%を除いた90%以内に入るのはどの値以上どの値以下か							
13	-2.8	0.007915452	0.00255513			以上		以下				
14	-2.6	0.013582969	0.004661188		下位2.5%と上位2.5%を除いた95%以内に入るのはどの値以上どの値以下か							
15	-2.4	0.02239453	0.008197536			以上		以下				
16	-2.2	0.035474593	0.013903448		下位0.5%と上位0.5%を除いた99%以内に入るのはどの値以上どの値以下か							
17	-2	0.053990967	0.022750132			以上		以下				
18	-1.8	0.078950158	0.035930319									

Sheet1

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

セル L2 には「=NORM.S.DIST(0,TRUE)」と入力されていて，0.5 が返されています．これにより，標準正規分布の累積分布関数の $z = 0$ における値が 0.5 であることが確認できます．

すなわち，逆関数については，累積分布関数の値 0.5 に対応する値は 0 になるので,「=NORM.S.INV(0.5)」は 0 になるはずです．

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

Point! **上側確率，下側確率，両側確率**

確率変数がある値 b 以上となる確率を，b についての上側確率といい，ある値 b 以下となる確率を，b についての下側確率といいます．

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

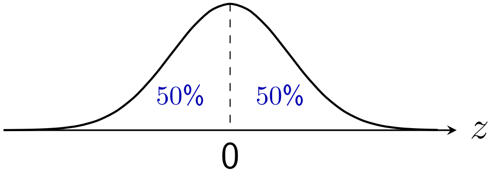
Point! 上側確率，下側確率，両側確率

また，正規分布のように左右対称な分布では，それに従う確率変数の中心からの距離がある値 b 以上（$b > 0$）となる確率を，b についての両側確率といいます．

標準正規分布では中心が 0 のため，「標準正規分布に従う確率変数の絶対値がある値 b 以上（$b > 0$）となる確率を，b についての両側確率という」と表現することができます．

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

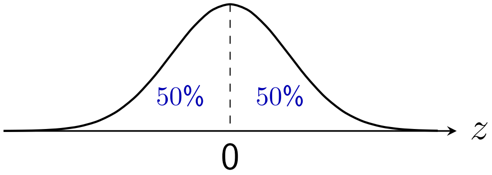
たとえば，標準正規分布は期待値 0 を中心として左右対称なので，それに従う確率変数 Z について，Z の値が 0 以上となる確率も，0 以下になる確率も，50%になります．



標準正規分布 $N(0,1^2)$ の確率密度関数のグラフ（$Z \geq 0$ となる確率も $Z \leq 0$ となる確率も 50%）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

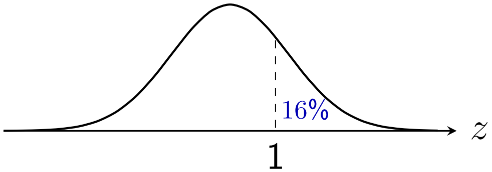
よって，標準正規分布に従う確率変数 Z において，0 についての上側確率も，下側確率も，50%になります．



標準正規分布 $N(0,1^2)$ の確率密度関数のグラフ（$Z \geq 0$ となる確率も $Z \leq 0$ となる確率も 50%）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

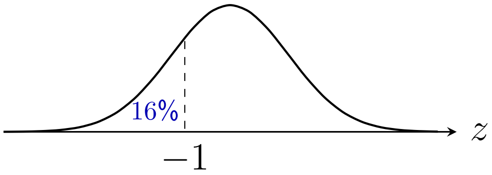
また，標準正規分布に従う確率変数 Z において，Z の値が 1 以上となる確率は 16%なので，1 についての上側確率は 16%になります．



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（$Z \geq 1$ となる確率は 16%）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

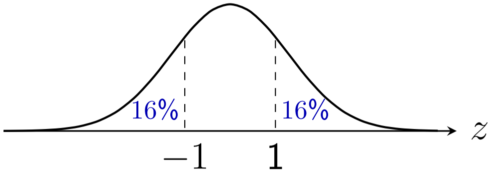
Z の値が -1 以下となる確率は 16%なので，-1 についての下側確率は 16%になります．なお，確率分布が左右対称であり，1 についての上側確率 16%であることからも，-1 についての下側確率は 16%であることがわかります．



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（$Z \leq -1$ となる確率は 16%）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

よって，Z の絶対値が 1 以上となる確率は $(16 \times 2)\%$，つまり，32% となるので，1 についての両側確率は 32% になります．



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（Z の絶対値が 1 以上となる確率は 32%）

例題6-4　標準正規分布の累積分布関数の逆関数の値を求める

Point!

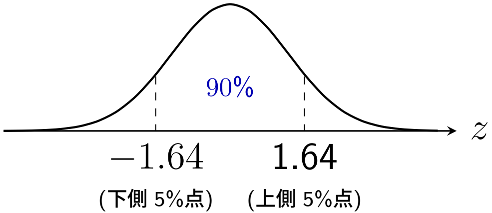
パーセント点

上側確率がc%であるような点（確率変数の値）のことを上側$\boldsymbol{c}$**%**点といい，下側確率がc%であるような点（確率変数の値）のことを下側$\boldsymbol{c}$**%**点といいます．

また，正規分布のように左右対称な分布では，両側確率がc%であるような点（確率変数の値）のことを両側$\boldsymbol{c}$**%**点といいます．

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

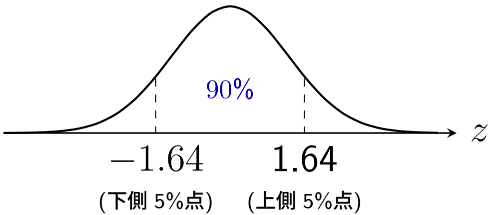
標準正規分布において，上位 5%以上に入るのは 1.64 以上の値なので，上側 5%点は 1.64 となります．



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（上側 5%点は 1.64）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

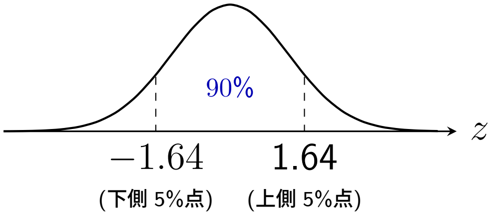
下位 5%以下に入るのは -1.64 以下の値なので，下側 5%点は -1.64 となります．なお，確率分布が左右対称であり，上側 5%点が 1.64 であることからも，下側 5%点は -1.64 であることがわかります．



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（下側 5%点は -1.64）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

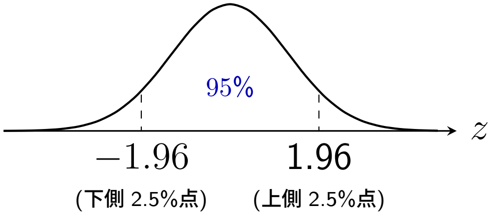
よって，両側 10%点は 1.64 となります.



標準正規分布 $N(0,1^2)$ の確率密度関数のグラフ（両側 10%点は 1.64）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

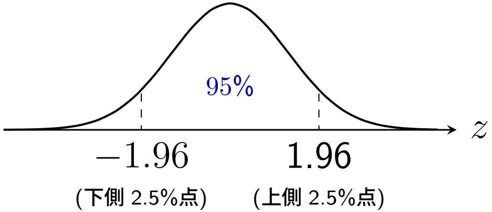
また，上位 2.5%以上に入るのは 1.96 以上の値なので，上側 2.5%点は 1.96 となります．



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（上側 2.5%点は 1.96）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

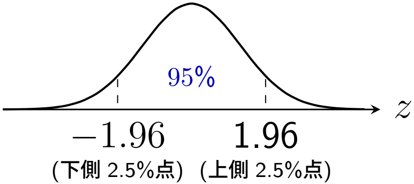
下位 2.5%以下に入るのは -1.96 以下の値なので，下側 2.5%点は -1.96 となります．なお，確率分布が左右対称であり，上側 2.5%点が 1.96 であることからも，下側 2.5%点は -1.96 であることがわかります．



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（下側 2.5%点は -1.96）

例題 6-4　標準正規分布の累積分布関数の逆関数の値を求める

よって，両側 5%点は 1.96 となります．



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（両側 5%点は 1.96）

例題 6-5　正規分布を標準化する

ある試験の点数 X が期待値 50，標準偏差 10 の正規分布 $N(50, 10^2)$ に従うとして，X を標準化しましょう．

また，標準化後の確率変数 Z の累積分布関数の値を求めましょう．

例題 6-5　正規分布を標準化する

Point! **確率変数の標準化**

期待値 μ，標準偏差 σ の確率変数 X から確率変数

$$Z = \frac{X-\mu}{\sigma} \qquad \left(Z = \frac{X - \text{期待値}}{\text{標準偏差}} \right)$$

への変換を標準化といい，変換された値を z 値といいます．

一般に，標準化された確率変数 Z の期待値は 0，標準偏差は 1 になります．X が正規分布 $N(\mu, \sigma^2)$ に従うときは，Z は標準正規分布 $N(0, 1^2)$ に従います．

例題 6-5　正規分布を標準化する

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	点数X	累積分布関数（X）		Z	累積分布関数（Z）		区間（X）	確率			区間（Z）	確率	
2	0	2.86652E-07		-5	2.86652E-07		X≦50	0.5			Z≦0	0.5	
3	1	4.79183E-07		-4.9	4.79183E-07		X≧50	0.5			Z≧0	0.5	
4	2	7.93328E-07		-4.8	7.93328E-07		X≦40	0.158655			Z≦-1	0.158655	
5	3	1.30081E-06		-4.7	1.30081E-06		X≧60	0.158655			Z≧1	0.158655	
6	4	2.11245E-06		-4.6	2.11245E-06		X≦30	0.02275			Z≦-2	0.02275	
7	5	3.39767E-06		-4.5	3.39767E-06		X≧70	0.02275			Z≧2	0.02275	
8	6	5.41254E-06		-4.4	5.41254E-06		X≦20	0.00135			Z≦-3	0.00135	
9	7	8.53991E-06		-4.3	8.53991E-06		X≧80	0.00135			Z≧3	0.00135	
10	8	1.33457E-05		-4.2	1.33457E-05		X≦10	3.17E-05			Z≦-4	3.17E-05	
11	9	2.06575E-05		-4.1	2.06575E-05		X≧90	3.17E-05			Z≧4	3.17E-05	
12	10	3.16712E-05		-4	3.16712E-05		40≦X≦60	0.682689	1σ区間		-1≦Z≦1	0.682689	1σ区間
13	11	4.80963E-05		-3.9	4.80963E-05		30≦X≦70	0.9545	2σ区間		-2≦Z≦2	0.9545	2σ区間
14	12	7.2348E-05		-3.8	7.2348E-05		20≦X≦80	0.9973	3σ区間		-3≦Z≦3	0.9973	3σ区間
15	13	0.0001078		-3.7	0.0001078		10≦X≦90	0.999937	4σ区間		-4≦Z≦4	0.999937	4σ区間
16	14	0.000159109		-3.6	0.000159109								
17	15	0.000232629		-3.5	0.000232629								

「X の累積分布関数の値」と「Z の累積分布関数の値」が一致することが確かめられます．

例題 6-5　正規分布を標準化する

また，例題 5-6 で求めた，正規分布 $N(50, 10^2)$ についての 1σ，2σ，3σ，4σ 区間と，例題 6-3 で求めた，標準正規分布についての 1σ，2σ，3σ，4σ 区間が，それぞれ標準化によって対応していることが確認できます．

たとえば，点数 40 の z 値は -1，点数 60 の z 値は 1 となり，1σ 区間の「$40 \leq X \leq 60$」と「$-1 \leq Z \leq 1$」は標準化によって対応しています．

例題 6-6　標準化前の値を求める

ある試験の点数 X が期待値 50，標準偏差 10 の正規分布 $N(50, 10^2)$ に従うとして，F 列の z 値を，標準化前の点数に戻したものを E 列に入力しましょう．

例題 6-6　標準化前の値を求める

ここで，標準化とは，期待値 μ，標準偏差 σ の確率変数 X から確率変数

$$Z = \frac{X - \mu}{\sigma} \qquad \left(Z = \frac{X - \text{期待値}}{\text{標準偏差}} \right)$$

への変換です．

例題 6-6　標準化前の値を求める

X を Z の式であらわすために，まず両辺に σ をかけると，

$$\sigma Z = X - \mu$$

となります．そして，これの両辺に μ を足すことにより

$$\boldsymbol{X = \sigma Z + \mu}$$

が得られます．

例題 6-6　標準化前の値を求める

	A	B	C	D	E	F	G	H	I	J	K	L
1	点数X	Z			点数X	Z		区間（X）	確率		区間（Z）	確率
2	0	-5		下位50%以下に入るのは何点以下か	50	0		X≦50	0.5		Z≦0	0.5
3	1	-4.9		上位50%以上に入るのは何点以上か	50	0		X≧50	0.5		Z≧0	0.5
4	2	-4.8		下位15.86553%以下に入るのは何点以下か	40	-1		X≦40	0.158655		Z≦-1	0.158655
5	3	-4.7		上位15.86553%以上に入るのは何点以上か	60	1		X≧60	0.158655		Z≧1	0.158655
6	4	-4.6		下位5%以下に入るのは何点以下か	33.55146	-1.64485						
7	5	-4.5		上位5%以上に入るのは何点以上か	66.44854	1.644854						
8	6	-4.4		下位2.5%以下に入るのはどの値以下か	30.40036	-1.95996						
9	7	-4.3		上位2.5%以上に入るのはどの値以上か	69.59964	1.959964						
10	8	-4.2		下位0.5%以下に入るのは何点以下か	24.24171	-2.57583						
11	9	-4.1		上位0.5%以上に入るのは何点以上か	75.75829	2.575829						
12	10	-4										
13	11	-3.9										
14	12	-3.8										
15	13	-3.7										
16	14	-3.6										
17	15	-3.5										

Sheet1

例題 6-7　正規分布の累積分布関数の逆関数の値を求める

例題 6-6 において，ある試験の点数 X が期待値 50，標準偏差 10 の正規分布 $N(50, 10^2)$ に従うとして，NORM.INV 関数を用いて，E 列に次をそれぞれ求めなおしてみましょう．そして，それが, (例題 6-6 で求めた)「F 列の z 値を標準化前の点数に戻したもの」と一致することを確認しましょう．

例題 6-7　正規分布の累積分布関数の逆関数の値を求める

下位 50%以下に入るのはどの値以下か，

上位 50%以上に入るのはどの値以上か，

下位 15.86553%以下に入るのはどの値以下か，

上位 15.86553%以上に入るのはどの値以上か，

例題 6-7　正規分布の累積分布関数の逆関数の値を求める

下位 5%以下に入るのはどの値以下か，
上位 5%以上に入るのはどの値以上か，
下位 2.5%以下に入るのはどの値以下か，
上位 2.5%以上に入るのはどの値以上か，
下位 0.5%以下に入るのはどの値以下か，
上位 0.5%以上に入るのはどの値以上か

例題6-7　正規分布の累積分布関数の逆関数の値を求める

	A	B	C	D	E	F	G	H	I	J	K	L
1	点数X	Z			点数X	Z		区間（X）	確率		区間（Z）	確率
2	0	-5		下位50%以下に入るのは何点以下か	50	0		X≦50	0.5		Z≦0	0.5
3	1	-4.9		上位50%以上に入るのは何点以上か	50	0		X≧50	0.5		Z≧0	0.5
4	2	-4.8		下位15.86553%以下に入るのは何点以下か	40	-1		X≦40	0.158655		Z≦-1	0.158655
5	3	-4.7		上位15.86553%以上に入るのは何点以上か	60	1		X≧60	0.158655		Z≧1	0.158655
6	4	-4.6		下位5%以下に入るのは何点以下か	33.55146	-1.64485						
7	5	-4.5		上位5%以上に入るのは何点以上か	66.44854	1.644854						
8	6	-4.4		下位2.5%以下に入るのはどの値以下か	30.40036	-1.95996						
9	7	-4.3		上位2.5%以上に入るのはどの値以上か	69.59964	1.959964						
10	8	-4.2		下位0.5%以下に入るのは何点以下か	24.24171	-2.57583						
11	9	-4.1		上位0.5%以上に入るのは何点以上か	75.75829	2.575829						
12	10	-4										
13	11	-3.9										
14	12	-3.8										
15	13	-3.7										
16	14	-3.6										
17	15	-3.5										

Sheet1

NORM.INV 関数で求めなおしても，まったく値が変わらないことがわかります．

例題 6-7　正規分布の累積分布関数の逆関数の値を求める

例題 5-5 で,「=NORM.DIST(50,50,10,TRUE)」と入力したら，0.5 が返されました．これにより，正規分布 $N(50, 10^2)$ の累積分布関数の $x = 50$ における値が 0.5 であることが確認できました．

すなわち，逆関数については，累積分布関数の値 0.5 に対応する値は 50 になるので,「=NORM.INV(0.5,50,10)」は 50 になるはずです．

第 7 章　点推定

7.1　母数と推定量

例題 7-1　母数とはなにかを理解する

問題 1-3 において，入力されているデータ（学生についての情報）について，点数の*母*平均と*母*分散と*母*標準偏差，および，標本 1 における点数の標本平均と標本分散と標本標準偏差を求めてみましょう．

例題 7-1　母数とはなにかを理解する

Point!

母平均，母分散，母標準偏差

母集団の平均値は母平均，母集団の分散は母分散，母集団の標準偏差は母標準偏差といわれます．このような，母集団の分布を特徴づける値（母平均など）のことを母数（パラメータ）といいます．

母数は多くの場合は未知であり，標本によりこれらを推定します．

例題 7-1　母数とはなにかを理解する

Point!

統計的推測

標本の性質をもとに，母集団の性質を調べることを統計的推測といいます．

統計的推測には，推定と（仮説）検定があります．

例題 7-1　母数とはなにかを理解する

Point!

統計的推測

推定は，標本を調べることにより母数を推測することです．推定には，点推定と区間推定という 2 つの方法があります．

点推定とは，平均値などを 1 つの値で推定することといい，区間推定とは，平均値などをある区間でもって推定することをいいます．

例題 7-1　母数とはなにかを理解する

Point!

統計的推測

一方，検定は，母集団の特徴についての仮説を立てて，その仮説が妥当かどうかを標本を調べて検証することです．

例題 7-1　母数とはなにかを理解する

Point!

標本平均，標本分散，標本標準偏差

母集団を表現する値を母数というのに対し，標本を表現する値を統計量といいます．統計量は標本から計算して求められるものです．

標本の平均値は標本平均，標本の分散は標本分散，標本の標準偏差は標本標準偏差といわれます．

例題 7-1　母数とはなにかを理解する

	A	B	C	D	E	F	G	H	I	J
1	母平均			49.6			標本平均			50.64
2	母分散			399.0247			標本分散			310.6304
3	母標準偏差			19.9756			標本標準偏差			17.62471
4							不偏分散			
5							不偏分散による標準偏差			
6							標準誤差			

例題 7-2　推定量とはなにかを理解する

問題 7-1 において，入力されているデータ（学生についての情報）について，標本 1 における点数の不偏分散，および，不偏分散による標準偏差を求めてみましょう．

例題 7-2　推定量とはなにかを理解する

Point! **不偏分散**

偏差の 2 乗和を「データの個数 -1」で割ったものは不偏分散とよばれます．

なお，標本分散は，偏差の 2 乗和を「データの個数」で割ったものです．

例題 7-2　推定量とはなにかを理解する

よって，標本分散に

$$\frac{\text{データの個数}}{\text{データの個数} - 1}$$

をかけると不偏分散が得られ，不偏分散に

$$\frac{\text{データの個数} - 1}{\text{データの個数}}$$

をかけると標本分散が得られるということになります.

例題 7-2　推定量とはなにかを理解する

Point!

推定量

標本から計算して求められる統計量が，母数を推定する目的で使われる場合は推定量とよばれます．

とくに，標本サイズを大きくするにつれて推定したい母数に限りなく近づくような推定量は一致推定量とよばれ，また，その期待値が推定したい母数と一致するような推定量は不偏推定量とよばれます．

例題 7-2　推定量とはなにかを理解する

一致推定量は，標本の大きさを増やしていったときに，推定量が母集団の真の値に限りなく近づいていく性質をもつ統計量です．つまり，十分に多くのデータを集めれば，母数を正しく推定できることを保証するものです．

標本平均は母平均の一致推定量であり，標本分散，不偏分散は母分散の一致推定量です．

例題 7-2　推定量とはなにかを理解する

また，不偏推定量は，標本を何度も取りなおして推定量を計算したときに，その平均値が母集団の真の値に一致する統計量です．つまり，平均的に偏りなく正しい推定を与える性質をもっています．

標本平均は母平均の不偏推定量です．標本分散は母分散を過少評価してしまい，母分散の不偏推定量ではありませんが，不偏分散は母分散の不偏推定量です．

例題 7-2　推定量とはなにかを理解する

一方，不偏分散の正の平方根は厳密には母標準偏差の不偏推定量ではありませんが，通常はこれで母標準偏差を推定します．ここでは，これを不偏分散による標準偏差とよびます．

Point!

不偏分散による標準偏差

$$\text{不偏分散による標準偏差} = \sqrt{\text{不偏分散}}$$

例題 7-2　推定量とはなにかを理解する

多くの場合，母集団のサイズは大きいことが多く，母平均や母分散を直接計算することはむずかしいため，標本を使って母集団の性質を推定します．

例題 7-2　推定量とはなにかを理解する

	A	B	C	D	E	F	G	H	I	J
1	母平均			49.6			標本平均			50.64
2	母分散			399.0247			標本分散			310.6304
3	母標準偏差			19.9756			標本標準偏差			17.62471
4							不偏分散			316.9698
5							不偏分散による標準偏差			17.80365
6							標準誤差			

同じデータから算出された標本分散と不偏分散については，不偏分散のほうが値が大きいことが確認できます．

例題 7-3　標準誤差を求める

問題 7-2 において，入力されているデータ（学生についての情報）について，標本 1 における点数の標準誤差を求めてみましょう．

例題 7-3　標準誤差を求める

標準誤差は，推定量がどれくらい正しいものかをあらわす指標です．

Point! **標準誤差**

標準誤差とは，標本平均のばらつきをあらわすものであり，次で求められます．

$$\text{標準誤差} = \sqrt{\frac{\text{不偏分散}}{\text{標本サイズ}}} = \frac{\text{不偏分散による標準偏差}}{\sqrt{\text{標本サイズ}}}$$

例題 7-3　標準誤差を求める

標本分散や不偏分散の算出に使っている標本平均は，どの標本から算出するかによってばらつきがあります．標準誤差は，このばらつきの大きさをあらわすものです．

例題 7-3　標準誤差を求める

不偏分散が同じなら，標本サイズが大きいほど，標準誤差は小さくなります．また，標本サイズが同じなら，標本のばらつきが大きいほど，標準誤差は大きくなります．

例題 7-3　標準誤差を求める

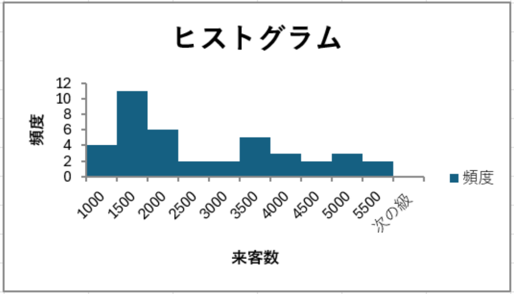
	A	B	C	D	E	F	G	H	I	J
1	母平均			49.6			標本平均			50.64
2	母分散			399.0247			標本分散			310.6304
3	母標準偏差			19.9756			標本標準偏差			17.62471
4							不偏分散			316.9698
5							不偏分散による標準偏差			17.80365
6							標準誤差			2.517816

7.2　分布の形状をあらわす基本統計量

例題 7-4　統計情報を求める

例題 2-5 において，入力されているデータ（ある市営プールにおける来客数の情報）における来客数について，Excel アドインのデータ分析ツールの「基本統計量」で求められる「統計情報」を求めてみましょう．

例題 7-4　統計情報を求める

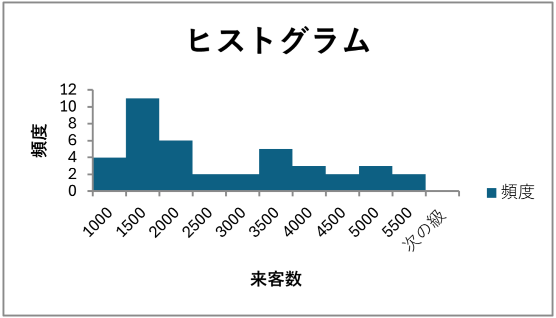


来客数	
平均	2528.85
標準誤差	215.0679
中央値 （メジアン）	1893.5
最頻値 （モード）	#N/A
標準偏差	1360.209
分散	1850169
尖度	-0.87191
歪度	0.674225
範囲	4560
最小	830
最大	5390
合計	101154
データの個数	40

例題 7-4　統計情報を求める

上記のような，平均，標準誤差，中央値，最頻値，標準偏差，分散，尖度，歪度，範囲，最小，最大，合計，データの個数のそれぞれの値の一覧が出てきます．これらの値を，例題 2-5 で作成したヒストグラムと一緒に見てみましょう．

例題 7-4　統計情報を求める



なお，ヒストグラムの各棒の高さは，左からそれぞれ,「1000 以下」,「1000 より大きく 1500 以下」,「1500 より大きく 2000 以下」, ⋯ ,「5000 より大きく 5500 以下」の区間の度数をあらわしています．

例題 7-4　統計情報を求める

平均について

平均値は「2528.85」であり,「2500 より大きく 3000 以下という区間」（左から 5 番目の区間）に属しています.

この場合，平均値は「もっとも度数が大きい区間」（1000 より大きく 1500 以下，左から 2 番目の区間）に属さず，それより右の区間に属していることがわかります.

例題 7-4　統計情報を求める

中央値について

中央値は「1893.5」であり,「1500 より大きく 2000 以下という区間」（左から 3 番目の区間）に属しています.

この場合，中央値は「もっとも度数が大きい区間」（1000 より大きく 1500 以下，左から 2 番目の区間）にも，平均値が属している「2500 より大きく 3000 以下という区間」（左から 5 番目の区間）にも，属していないということになります.

例題 7-4　統計情報を求める

最頻値について

最頻値は「#N/A」と表示され，存在しないことが確認できます．最頻値を算出する際，まったく同じ値にならないと同じと判断されませんが，このデータのように相異なる値が豊富にあるときは，まったく同じ値であることはまれです．

このような場合，ヒストグラムのようにデータを区切ることにより，「もっとも度数が大きい区間」（1000 より大きく 1500 以下，左から 2 番目の区間）に着目します．ただし，区間の区切り方により度数分布が異なるということには注意しましょう．

例題 7-4　統計情報を求める

標準偏差について

「不偏分散による標準偏差」は約「1360.2」であることがわかります．

Excel アドインのデータ分析ツールの「基本統計量」で求められる「標準偏差」は，STDEV.P 関数で求められる「標準偏差」ではなく，STDEV.S 関数で求められる「不偏分散による標準偏差」であることに注意しましょう．

分散について

不偏分散は約「1850169」であることがわかります．

Excel アドインのデータ分析ツールの「基本統計量」で求められる「分散」は，VAR.P 関数で求められる「標本分散」ではなく，VAR.S 関数で求められる「不偏分散」であることに注意しましょう．

例題 7-4　統計情報を求める

つぎに，尖度と歪度は，次のような，分布の形状をあらわす単位をもたない数値です．

Point!

尖度と歪度

尖度は分布の形がどれくらいとがって（尖って）いるのかをあらわします．正規分布の尖度は 0 です．

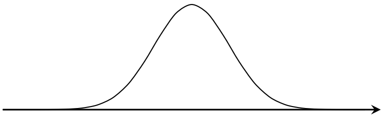
- 尖度が大きいほど,「山がとがっている」分布になります．
- 尖度が小さいほど,「山がつぶれている」分布になります．

例題 7-4　統計情報を求める

Point! 尖度と歪度

歪度は分布の形がどれくらいゆがんで（歪んで）いるのかをあらわします．正規分布のような左右対称の分布の歪度は 0 です．

- 歪度が正のとき,「山が左にかたよっている」分布になります．
- 歪度が負のとき,「山が右にかたよっている」分布になります．



正規分布（尖度 0，歪度 0）

例題 7-4　統計情報を求める

尖度について

尖度は約「-0.87」であり 0 より小さいので，正規分布よりとがっていないような分布をもつことがわかります．

例題 7-4　統計情報を求める

歪度について

歪度は約「0.67」であり 0 より大きいので，山が左にかたよっているような形状の分布をもつことがわかります．

例題 7-4　統計情報を求める

そして，範囲とは，次のような，最大値と最小値の差のことをいいます．

Point! **範囲**

データの範囲の大きさ，つまり，

$$\text{最大値} - \text{最小値}$$

を範囲（レンジ）といいます．範囲の大きさは，データの収まる範囲の広さを示しています．

例題 7-4　統計情報を求める

範囲を求めるときは最大値と最小値しか使わないので，範囲からデータの個数がわかるわけでもないし，平均値など全体の傾向がわかるわけでもありません．

極端に大きい値や小さい値に影響されやすいことも意識しておきましょう．

例題 7-4　統計情報を求める

範囲について

最大値は「5390」，最小値は「830」なので，範囲は

$$5390 - 830 = 4560$$

と計算され,「4560」であることがわかります.

第8章　中心極限定理

8.1　大数の法則

例題 8-1　標本平均を調べて母平均と比較する

サイコロを 4 回振るという試行を 100 回くり返し，それぞれの試行について，出た目の平均値を求めましょう．

そして，平均値のグラフを作成し，サイコロを振ったときに出る目 X の期待値 3.5 と比較してみましょう．

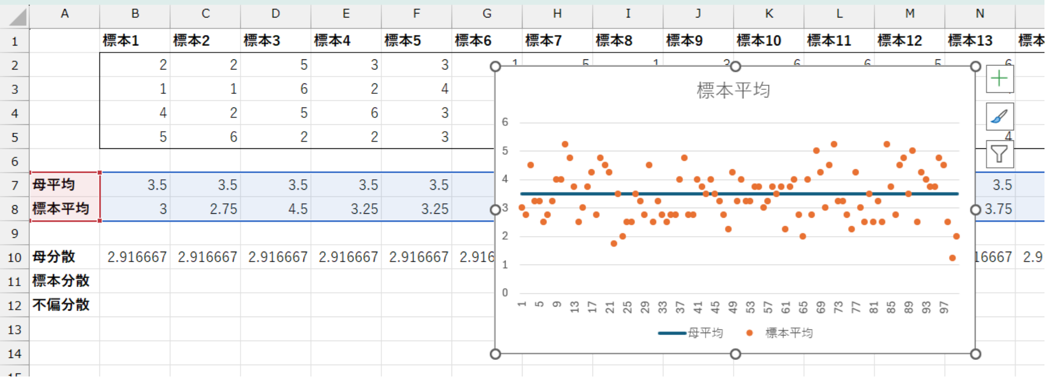
例題 8-1　標本平均を調べて母平均と比較する

この例題では，サイコロを振ったときに出る目 X が従う確率分布をとるような母集団を想定しています（X の期待値は 3.5，分散は $35/12$）．母平均は 3.5，母分散は $35/12$ ということになります．

x	1	2	3	4	5	6
$P(X=x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

このような母集団から大きさ 4 の標本を 100 回無作為抽出します．それぞれの標本について，標本平均を求めて，グラフを作成することにより母平均 3.5 と比較してみましょう．

例題 8-1　標本平均を調べて母平均と比較する



	A	B	C	D	E	F
1		標本1	標本2	標本3	標本4	標本5
2		2	2	5	3	3
3		1	1	6	2	4
4		4	2	5	6	3
5		5	6	2	2	3
6						
7	母平均	3.5	3.5	3.5	3.5	3.5
8	標本平均	3	2.75	4.5	3.25	3.25
9						
10	母分散	2.916667	2.916667	2.916667	2.916667	2.916667
11	標本分散					
12	不偏分散					

標本平均は母平均の付近に集まっていることが確かめられます.

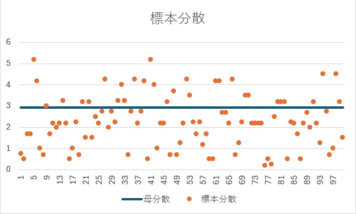
例題 8-2　標本分散と不偏分散を調べて母分散と比較する

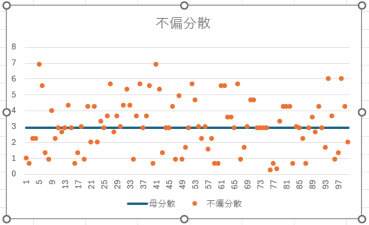
例題 8-1 において，100 回くり返された「サイコロを 4 回振る」というそれぞれの試行について，出た目の標本分散と不偏分散を求めましょう．

そして，標本分散と不偏分散について，それぞれのグラフを作成し，サイコロを振ったときに出る目 X の分散 $35/12$ と比較してみましょう．

例題8-2　標本分散と不偏分散を調べて母分散と比較する

標本平均	3.5	4	3.25	2.75	3.25	2.75	3	3.25	4	4.25	3.75	3	2.75	3.5
母分散	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667
標本分散	0.75	0.5	1.6875	1.6875	5.1875	4.1875	1	0.6875	3	1.6875	2.1875	2	2.1875	3.25
不偏分散	1	0.666667	2.25	2.25	6.916667	5.583333	1.333333	0.916667	4	2.25	2.916667	2.666667	2.916667	4.333333





不偏分散は母分散の付近に集まっていることが確かめられます．
一方，標本分散は，母分散と比較すると全体的に小さくなっていることがわかります．

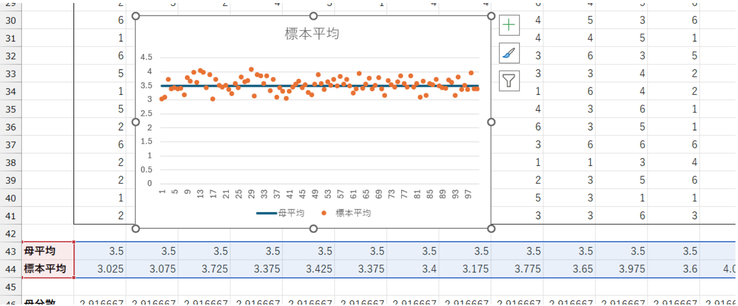
例題 8-3　標本サイズを大きくして標本平均を調べて母平均と比較する

サイコロを 40 回振るという試行を 100 回くり返し，それぞれの試行について，出た目の平均値を求めましょう．

そして，平均値のグラフを作成し，サイコロを振ったときに出る目 X の期待値 3.5 と比較してみましょう．

また，例題 8-1 における標本平均のグラフと比べてみましょう．

例題8-3　標本サイズを大きくして標本平均を調べて母平均と比較する



標本平均は母平均の付近に集まっていることが確かめられます．
また，例題8-1における標本平均のグラフと比べると，さらに母平均の付近に集まっていることがわかります．

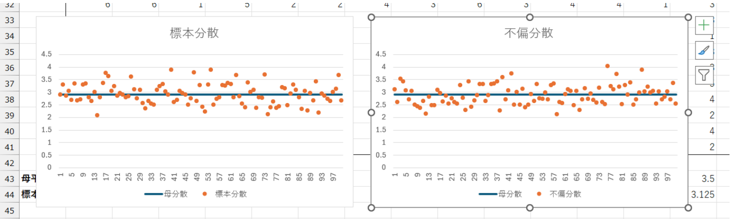
例題 8-4　標本サイズを大きくして標本分散と不偏分散を調べて母分散と比較する

例題 8-3 において，100 回くり返された「サイコロを 40 回振る」というそれぞれの試行について，出た目の標本分散と不偏分散を求めましょう．

そして，標本分散と不偏分散について，それぞれのグラフを作成し，サイコロを振ったときに出る目 X の分散 $35/12$ と比較してみましょう．

また，例題 8-2 における標本分散のグラフ，不偏分散のグラフとそれぞれ比べてみましょう．

例題 8-4　標本サイズを大きくして標本分散と不偏分散を調べて母分散と比較する



母分散	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.916667	2.
標本分散	3.034375	2.54	3.444375	3.34	2.999375	2.634375	2.9775	2.444375	2.384375	2.309375	2.5775	2.099375	2.74	2.409375	
不偏分散	3.112179	2.605128	3.532692	3.425641	3.076282	2.701923	3.053846	2.507051	2.445513	2.36859	2.64359	2.153205	2.810256	2.471154	2.

不偏分散，標本分散ともに母分散の付近に集まっていることが確かめられます．
また，例題 8-2 における標本分散のグラフ，不偏分散のグラフと比べると，さらに母分散の付近に集まっていることがわかります．

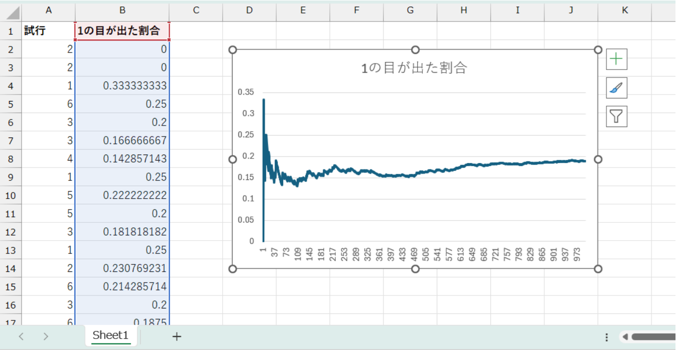
例題 8-5　大数の法則を確認する

サイコロを 1000 回以上振り，各回について，それまでに 1 の目が出た割合をそれぞれ求めましょう．

そして，各回までの 1 の目が出た割合の折れ線グラフを作成し,「1 の目が出る確率 $1/6$」と比較してみます．

サイコロを何回も振ったら，1 の目が出る割合が $1/6$ に近づくかどうか，確かめてみましょう．

例題 8-5　大数の法則を確認する



グラフを見ると，サイコロを振る回数が増えるにつれて，1 の目が出た割合が「1 の目が出る確率 $1/6$ $(= 0.166\cdots)$」に収束していくのがわかります．

例題 8-5　大数の法則を確認する

このように，サイコロを振る回数を増やすと各目の出る割合は $1/6$ に近づくので，出る目の平均値（標本平均）は

$$1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3.5$$

に近づく，つまり，サイコロを振ったときに出る目 X の期待値 3.5（X が従う確率分布をとるような母集団の平均値（母平均））に近づくといえます．

例題 8-5　大数の法則を確認する

一般に，確率 p で起こる事象において，試行回数を増やすと，その事象が実際に起こる割合は p に近づきます．このことは，大数の法則によって保証されます．

例題 8-5　大数の法則を確認する

Point!

大数の法則

ある母集団から標本を無作為抽出をする場合，標本サイズを大きくするにつれて，標本平均は母平均に限りなく近づきます．

この現象が成り立つことを保証する確率論の定理は，大数の法則とよばれています．

大数の法則が成り立つため，標本平均は母平均の一致推定量になります．

8.2　中心極限定理

例題 8-6　中心極限定理を確認する

サイコロを 2 回振るという試行を 1000 回くり返し，それぞれの試行について，出た目の平均値（標本平均）を求めましょう．
そして，標本平均の分布が正規分布に似ていることを確かめましょう．

さらに，サイコロを 5 回振るという試行を 1000 回くり返し，それぞれの試行について，出た目の平均値（標本平均）を求めましょう．
そして，標本平均の分布が正規分布にもっと似ていることを確かめましょう．

例題 8-6　標本平均の分布を調べる

Point!

中心極限定理

平均値 μ，分散 σ^2 である母集団から，大きさ n の標本の無作為抽出をくり返すとき,(考えられるすべての標本についての）標本平均の分布は，n が大きくなるにつれて正規分布 $N(\mu, \sigma^2/n)$ に近づきます．

また，その標本に含まれるデータの合計の分布は，n が大きくなるにつれて正規分布 $N(n\mu, n\sigma^2)$ に近づきます．

これらの現象が成り立つことを保証する確率論の定理は，中心極限定理とよばれています．

例題 8-6　標本平均の分布を調べる

標本サイズが大きいほど，元がどんな分布でも，そこから抽出された標本平均の分布は，正規分布に近づきます．

そして，標本サイズが大きいほど,「標本平均の平均値」は母平均に近づき,「標本平均の分散」は「母分散/標本サイズ」に近づきます．

例題 8-6　標本平均の分布を調べる

これは，標本サイズが大きいほど，その標本平均のばらつきが小さくなり，標本平均が母平均のより近くに集まる，つまり，母平均をより正確に推測できるようになることを意味しています．

例題 8-6　標本平均の分布を調べる

なお，母集団が正規分布 $N(\mu, \sigma^2)$ をとる場合は，標本サイズ n によらず，標本平均の分布は正規分布 $N(\mu, \sigma^2/n)$ になります（正規分布の再生性）．

例題 8-6　標本平均の分布を調べる

この例題では，サイコロを振ったときに出る目 X が従う確率分布をとるような母集団を想定しています（X の期待値は 3.5，分散は $35/12$）．

x	1	2	3	4	5	6
$P(X=x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

例題 8-6　標本平均の分布を調べる

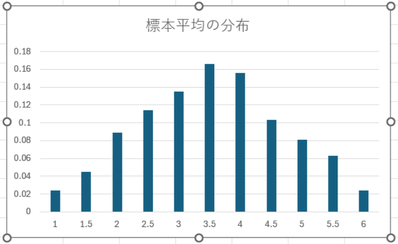
このような母集団から，大きさ 2 の標本を 1000 回無作為抽出し，標本平均の分布を調べてみます．

さらに，大きさ 5 の標本を 1000 回無作為抽出し，標本平均の分布を調べてみます．

このような分布は，中心極限定理により，標本の大きさ n を大きくすると，期待値 3.5，分散 $((35/12)/n =)$ $35/12n$ の正規分布 $N(3.5, 35/12n)$ に近づくはずです．

例題 8-6　標本平均の分布を調べる

	A	B	C	D	E	F	G
1		X_1	X_2	標本平均		標本平均	相対度数
2	標本1	6	5	5.5		1	0.024
3	標本2	4	1	2.5		1.5	0.045
4	標本3	1	2	1.5		2	0.089
5	標本4	2	6	4		2.5	0.114
6	標本5	5	1	3		3	0.135
7	標本6	4	1	2.5		3.5	0.166
8	標本7	5	2	3.5		4	0.156
9	標本8	6	4	5		4.5	0.103
10	標本9	4	1	2.5		5	0.081
11	標本10	2	3	2.5		5.5	0.063
12	標本11	3	3	3		6	0.024
13	標本12	3	3	3			
14	標本13	6	2	4			
15	標本14	6	4	5			
16	標本15	3	6	4.5			
17	標本16	6	5	5.5			

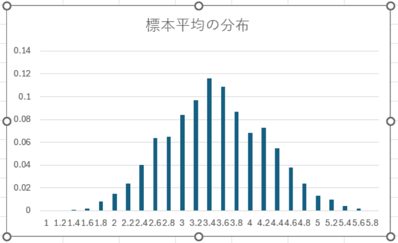


Sheet1

大きさ 2 の標本を 1000 回無作為抽出したときの標本平均の分布は，正規分布に似ていることが確かめられます．

例題 8-6　標本平均の分布を調べる

	A	B	C	D	E	F	G	H	I	J
1		X_1	X_2	X_3	X_4	X_5	標本平均		標本平均	相対度数
2	標本1	4	3	2	5	4	3.6		1	0
3	標本2	2	4	5	4	3	3.6		1.2	0
4	標本3	2	3	6	2	4	3.4		1.4	0.001
5	標本4	1	3	1	6	2	2.6		1.6	0.002
6	標本5	6	5	5	5	5	5.2		1.8	0.008
7	標本6	6	3	6	3	1	3.8		2	0.015
8	標本7	3	5	2	1	3	2.8		2.2	0.024
9	標本8	6	1	4	3	4	3.6		2.4	0.04
10	標本9	4	6	5	6	2	4.6		2.6	0.064
11	標本10	3	2	1	4	2	2.4		2.8	0.065
12	標本11	3	3	5	3	2	3.2		3	0.084
13	標本12	4	2	4	6	6	4.4		3.2	0.097
14	標本13	4	5	1	4	2	3.2		3.4	0.116
15	標本14	5	2	1	3	6	3.4		3.6	0.109
16	標本15	3	3	2	4	2	2.8		3.8	0.087
17	標本16	2	2	4	2	3	2.6		4	0.068



Sheet1

大きさ 5 の標本を 1000 回無作為抽出したときの標本平均の分布は，正規分布にもっと似ていることが確かめられます．

例題 8-7　中心極限定理を確認する

例題 8-6(1) では，サイコロを 2 回振るという試行を 1000 回くり返し，それぞれの試行について，出た目の平均値（標本平均）を求めました．そして，標本平均について，相対度数を計算し，分布を調べました．

また，例題 8-6(2) では，サイコロを 5 回振るという試行を 1000 回くり返し，同様のことを調べました．

このような分布は，中心極限定理により，標本の大きさ n を大きくすると，期待値 3.5，分散 $35/12n$ の正規分布 $N(3.5, 35/12n)$ に近づくはずです．

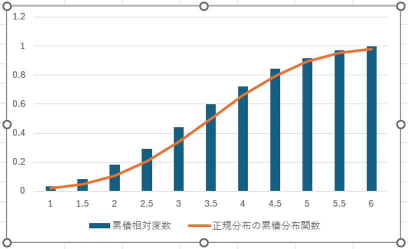
例題 8-7　中心極限定理を確認する

ここでは，例題 8-6(1) において，標本平均の累積相対度数のグラフが，正規分布 $N(3.5, 35/(12 \times 2))$ の累積分布関数のグラフに似ていることを確かめましょう．

さらに，例題 8-6(2) において，標本平均の累積相対度数のグラフが，正規分布 $N(3.5, 35/(12 \times 5))$ の累積分布関数のグラフにとても似ていることも確かめましょう．

例題 8-7　中心極限定理を確認する

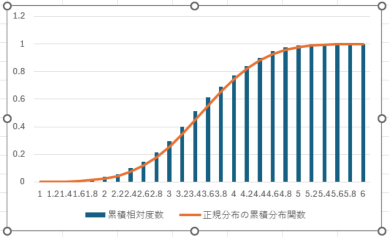
	X_2	標本平均	標本平均	相対度数	累積相対度数	正規分布の累積分布関数
6	5	5.5	1	0.033	0.033	0.019216965
4	6	5	1.5	0.051	0.084	0.04884498
3	3	3	2	0.098	0.182	0.107096513
1	5	3	2.5	0.109	0.291	0.203812974
6	4	5	3	0.15	0.441	0.33942265
4	5	4.5	3.5	0.16	0.601	0.5
4	2	3	4	0.122	0.723	0.66057735
3	3	3	4.5	0.122	0.845	0.796187026
3	5	4	5	0.069	0.914	0.892903487
1	5	3	5.5	0.056	0.97	0.95115502
5	4	4.5	6	0.03	1	0.980783035
4	2	3				
6	3	4.5				
4	6	5				
2	2	2				
6	2	4				



「大きさ 2 の標本を 1000 回無作為抽出したときの標本平均の累積相対度数のグラフ」が,「正規分布 $N(3.5, 35/(12 \times 2))$ の累積分布関数のグラフ」に似ていることが確かめられます．

例題 8-7　中心極限定理を確認する

	X_4	X_5	標本平均		標本平均	相対度数	累積相対度数	正規分布の累積分布関数
6	1	6	4.2		1	0	0	0.000531557
1	2	6	3.2		1.2	0.001	0.001	0.001300201
4	3	5	4		1.4	0.003	0.004	0.0029839
3	5	2	3.8		1.6	0.005	0.009	0.006428897
1	4	4	3.4		1.8	0.01	0.019	0.013013037
1	4	5	3.4		2	0.02	0.039	0.024767307
2	2	5	3.2		2.2	0.018	0.057	0.044368355
5	3	3	3.2		2.4	0.044	0.101	0.074900027
4	1	5	3.6		2.6	0.045	0.146	0.11932325
1	5	6	3		2.8	0.067	0.213	0.179698385
3	2	3	2.8		3	0.084	0.297	0.25634538
5	1	2	3.2		3.2	0.102	0.399	0.347236482
1	2	3	2.4		3.4	0.113	0.512	0.447915051
5	6	4	5		3.6	0.102	0.614	0.552084949
5	3	6	3.4		3.8	0.078	0.692	0.652763518
6	2	4	4.4		4	0.08	0.772	0.74365462



「大きさ 5 の標本を 1000 回無作為抽出したときの標本平均の累積相対度数のグラフ」が,「正規分布 $N(3.5, 35/(12 \times 5))$ の累積分布関数のグラフ」にとても似ていることが確かめられます.

第 9 章　母平均の区間推定（母分散既知）

9.1　母平均の 95%信頼区間（母分散既知）

例題 9-1　母平均の 95%信頼区間を求める（母分散既知）

ある試験の点数のデータを母集団とし，そこから無作為に抽出された大きさ 100 の標本により，母平均 μ の 95%信頼区間を求めましょう．ただし，母集団は正規分布をとるとし，母分散 σ^2 は 400 であることがわかっているとします．

例題 9-1　母平均の 95%信頼区間を求める（母分散既知）

Point!

区間推定

標本から得られた値を使って，ある区間でもって母平均などの母数を推定する方法は区間推定とよばれます．このときの区間のことを信頼区間といいます．

例題 9-1　母平均の 95%信頼区間を求める（母分散既知）

Point!

区間推定

母平均の区間推定では**95%**信頼区間を求めることが多いですが，この「95%」という割合は信頼係数といわれます．これは,「母集団から標本を取ってきて信頼区間を求めるという作業を 100 回くり返したときに，95 回くらいはその区間のなかに母平均が含まれる」ということを意味します．

例題 9-1　母平均の 95%信頼区間を求める（母分散既知）

Point! 母平均の 95%信頼区間（母分散既知）

正規分布 $N(\mu, \sigma^2)$ をとる母集団について，母分散 σ^2 はわかっているとし，母平均 μ を推定することを考えます．

この母集団から大きさ n の標本を無作為抽出するとき，標本平均を $\overline{X}$ とすると，次の式から母平均 μ の 95%信頼区間を求めることができます．

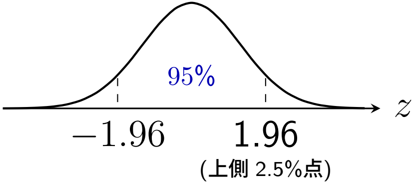
例題 9-1　母平均の 95%信頼区間を求める（母分散既知）

Point!

母平均の 95%信頼区間（母分散既知）

$$\overline{X} - \text{「標準正規分布における上側 2.5\%点」} \times \sqrt{\frac{\sigma^2}{n}} \leq \mu$$

$$\leq \overline{X} + \text{「標準正規分布における上側 2.5\%点」} \times \sqrt{\frac{\sigma^2}{n}}$$



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（上側 2.5%点は 1.96）

例題 9-1　母平均の 95%信頼区間を求める（母分散既知）

母平均の 95%信頼区間（母分散既知）

ここで，標準正規分布における上側 2.5%点は約 1.96 なので，上の式は次のようになります．

$$\overline{X} - 1.96 \times \sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \overline{X} + 1.96 \times \sqrt{\frac{\sigma^2}{n}}$$

なお，上の式のなかの $\sqrt{\frac{\sigma^2}{n}}$ は,「母標準偏差/$\sqrt{\text{標本サイズ}}$」であり，「標本平均の標準偏差」になります．

例題 9-1　母平均の 95%信頼区間を求める（母分散既知）

母平均 μ の 95%信頼区間はおおよそ

$$46.1 \leq \mu \leq 54.0$$

となります.

	A	B	C	D	E	F	G
1	**母分散 σ^2=400**		**標本平均**	50.04			
2							
3	標本		**標準正規分布における上側2.5%点**			1.959963985	
4	72						
5	50		**母平均 μ の95%信頼区間**		**46.12007**	**$\leq$ 母平均 μ $\leq$**	**53.95993**
6	48						
7	50						
8	68						
9	48						
10	44						

例題 9-1　母平均の 95%信頼区間を求める（母分散既知）

なお，母集団が正規分布をとるとは限らない場合でも，標本サイズが十分大きければ，中心極限定理によって標本平均の分布は正規分布に近似すると考えて，区間推定をすることがあります．

例題 9-2　母平均の 95%信頼区間を求めてグラフを作成する（母分散既知）

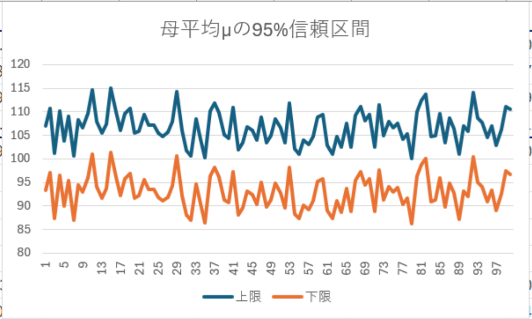
正規分布 $N(100, 7^2)$ をとるような母集団を想定します．母分散 49 はわかっているとし，母平均 μ を推定することを考えます．このような母集団から大きさ 4 の標本を 100 回無作為抽出しましょう．

そして，それぞれの標本から算出される母平均の 95%信頼区間の上限と下限についての折れ線グラフを作成しましょう．

また，算出された各 95%信頼区間のなかに，母平均 100 が含まれるなら「○」，含まれないなら「×」が表示されるようにし,「○」の割合を求めてみましょう．

例題 9-2　母平均の 95%信頼区間を求めてグラフを作成する（母分散既知）

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	母分散 σ^2=49												
2		100.2411	101.6659	90.91462	110.3907	97.							4.0695
3		93.53686	101.2463	102.7709	100.1729	88.3							.38202
4		112.6701	109.4919	91.92676	95.6686	90.9							9.0702
5		94.08363	103.4388	91.62082	107.1082	111.							03.392
6	標本平均	100.1329	103.9607	94.30828	103.3351	96.9							0.9784
7													
8	標準正規分布における上側2.5%点				1.959964								
9													
10	母平均 μ の95%信頼区間												
11	上限	106.9928	110.8206	101.1682	110.195	10							7.8383
12	下限	93.27306	97.10086	87.44841	96.47524	90.0							.11857
13		○	○	○	○	○	○	○	○	○	○	×	○
14													
15	成功率	95%											



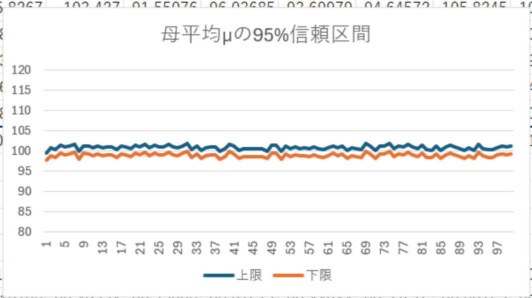
例題 9-3　標本サイズを大きくして母平均の 95%信頼区間を求めてグラフを作成する（母分散既知）

例題 9-2 における作業を，各標本サイズを 200 にしてやってみるとどうなるか，試してみましょう．

そして，例題 9-2 で作成した「大きさ 4 の標本から算出される母平均の 95%信頼区間の上限と下限についての折れ線グラフ」と，ここで作成した「大きさ 200 の標本から算出される母平均の 95%信頼区間の上限と下限についての折れ線グラフ」を比較してみましょう．

例題 9-3　標本サイズを大きくして母平均の 95%信頼区間を求めてグラフを作成する（母分散既知）

196		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
197		97.52828	110.6029	96.3322	102.266	105.8267	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	102.5443
198		93.91491	96.30705	105.652	96.66452	95.8							05.0232
199		104.8708	99.73787	90.03201	100.7488	98.3							9.22892
200		91.3396	103.6215	89.54616	104.4692	94.6							4.43691
201		83.72198	94.37484	88.98287	103.056	90.8							01.4501
202	標本平均	98.67787	99.93119	99.46839	100.4799	100							100.256
203													
204	標準正規分布における上側2.5%点			1.959964									
205													
206	母平均μの95%信頼区間												
207	上限	99.64801	100.9013	100.4385	101.45	101							01.2261
208	下限	97.70774	98.96105	98.49825	99.50977	99.12198	99.40334	99.73998	98.01617	99.44044	99.37615	98.89163	99.28582
209		×	○	○	○	○	○	○	×	○	○	○	○
210													
211	成功率	95%											
212													



例題 9-3　標本サイズを大きくして母平均の 95%信頼区間を求めてグラフを作成する（母分散既知）

ここで作成した「大きさ 200 の標本から算出される母平均の 95%信頼区間の上限と下限についての折れ線グラフ」は，例題 9-2 で作成した「大きさ 4 の標本から算出される母平均の 95%信頼区間の上限と下限についての折れ線グラフ」と比べると，区間の幅が小さく，ばらつきも小さいことが確認できます．

一般に，標本サイズが大きいほど，その標本から母分散を使って算出される母平均の信頼区間の幅は小さくなります．

9.2　母平均のさまざまな信頼区間（母分散既知）

例題 9-4　母平均の 90%信頼区間を求めてグラフを作成する（母分散既知）

例題 9-2 において，母平均 μ の 95%信頼区間を，90%信頼区間に変更しましょう．

そして，例題 9-2 で作成した「母平均の 95%信頼区間の上限と下限についての折れ線グラフ」と，ここで作成した「母平均の 90%信頼区間の上限と下限についての折れ線グラフ」を比較してみましょう．

例題 9-4　母平均の 90%信頼区間を求めてグラフを作成する（母分散既知）

母平均の 90%信頼区間（母分散既知）

正規分布 $N(\mu, \sigma^2)$ をとる母集団について，母分散 σ^2 はわかっているとし，母平均 μ を推定することを考えます．

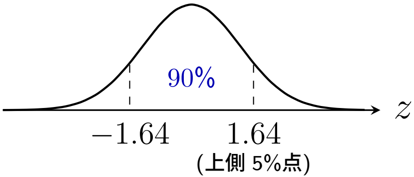
この母集団から大きさ n の標本を無作為抽出するとき，標本平均を $\overline{X}$ とすると，次の式から母平均 μ の 90%信頼区間を求めることができます．

例題 9-4　母平均の 90%信頼区間を求めてグラフを作成する（母分散既知）

母平均の 90%信頼区間（母分散既知）

$$\overline{X} - \text{「標準正規分布における上側 5\%点」} \times \sqrt{\frac{\sigma^2}{n}} \leq \mu$$

$$\leq \overline{X} + \text{「標準正規分布における上側 5\%点」} \times \sqrt{\frac{\sigma^2}{n}}$$



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（上側 5%点は 1.64）

例題 9-4　母平均の 90%信頼区間を求めてグラフを作成する（母分散既知）

Point!

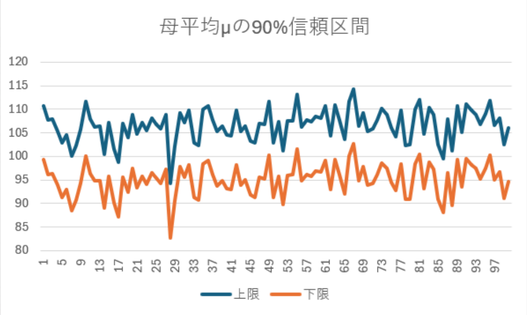
母平均の 90%信頼区間（母分散既知）

ここで，標準正規分布における上側 5%点は約 1.64 なので，上の式は次のようになります．

$$\overline{X} - 1.64 \times \sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \overline{X} + 1.64 \times \sqrt{\frac{\sigma^2}{n}}$$

例題 9-4　母平均の 90%信頼区間を求めてグラフを作成する（母分散既知）

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	母分散 σ^2=49												
2		104.1905	103.3445	101.7057	102.1105	94.3							3.0346
3		109.494	105.6085	107.4802	93.34374	101.							3.9449
4		96.67193	105.7865	92.70198	95.43589	98.2							2.4408
5		109.7011	93.06456	106.6003	107.585	94.7							.84605
6	標本平均	105.0144	101.951	102.122	99.61879	97.1							0.5666
7													
8	標準正規分布における上側5%点				1.644854								
9													
10	母平均 μ の90%信頼区間												
11	上限	110.7714	107.708	107.879	105.3758	102.							6.3236
12	下限	99.2574	96.19404	96.36506	93.8618	91.3							4.8096
13		○	○	○	○	○	○	○	○	○	×	○	○
14													
15	成功率	90%											



例題 9-4　母平均の 90%信頼区間を求めてグラフを作成する（母分散既知）

ここで作成した「母平均の 90%信頼区間の上限と下限についての折れ線グラフ」は，例題 9-2 で作成した「母平均の 95%信頼区間の上限と下限についての折れ線グラフ」と比べると，区間の幅が小さいことが確認できます．

例題 9-5　母平均の 99%信頼区間を求めてグラフを作成する（母分散既知）

例題 9-2 において，母平均 μ の 95%信頼区間を，99%信頼区間に変更しましょう．

そして，例題 9-2 で作成した「母平均の 95%信頼区間の上限と下限についての折れ線グラフ」と，ここで作成した「母平均の 99%信頼区間の上限と下限についての折れ線グラフ」を比較してみましょう．

例題 9-5　母平均の 99%信頼区間を求めてグラフを作成する（母分散既知）

Point!

母平均の 99%信頼区間（母分散既知）

正規分布 $N(\mu, \sigma^2)$ をとる母集団について，母分散 σ^2 はわかっているとし，母平均 μ を推定することを考えます．

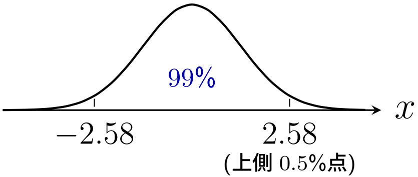
この母集団から大きさ n の標本を無作為抽出するとき，標本平均を $\overline{X}$ とすると，次の式から母平均 μ の 99%信頼区間を求めることができます．

例題 9-5　母平均の 99%信頼区間を求めてグラフを作成する（母分散既知）

Point! 母平均の 99%信頼区間（母分散既知）

$$\overline{X} - \text{「標準正規分布における上側 0.5\%点」} \times \sqrt{\frac{\sigma^2}{n}} \leq \mu$$

$$\leq \overline{X} + \text{「標準正規分布における上側 0.5\%点」} \times \sqrt{\frac{\sigma^2}{n}}$$



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（上側 0.5%点は 2.58）

例題 9-5　母平均の 99%信頼区間を求めてグラフを作成する（母分散既知）

Point!

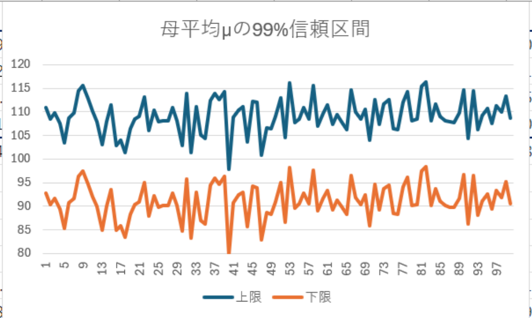
母平均の 99%信頼区間（母分散既知）

ここで，標準正規分布における上側 0.5%点は約 2.58 なので，上の式は次のようになります．

$$\overline{X} - 2.58 \times \sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \overline{X} + 2.58 \times \sqrt{\frac{\sigma^2}{n}}$$

例題 9-5　母平均の 99%信頼区間を求めてグラフを作成する（母分散既知）

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	母分散 σ^2=49												
2		111.9531	104.4364	107.7569	99.18275	94.9							)6.0007
3		97.27271	97.05033	94.35405	92.96326	90.2							.00.719
4		97.18882	92.08396	103.35	100.2833	100.							;.86311
5		101.0392	104.1266	97.35181	101.5498	92.1							)3.2158
6	標本平均	101.8635	99.42433	100.7032	98.49477	94.4							.94964
7													
8	標準正規分布における上側0.5%点				2.575829								
9													
10	母平均 μ の99%信頼区間												
11	上限	110.8789	108.4397	109.7186	107.5102	103.							.07.965
12	下限	92.84806	90.40892	91.68779	89.47936	85.3							.93424
13		○	○	○	○	○	○	○	○	○	○	○	○
14													
15	成功率	99%											
16													



例題 9-5　母平均の 99%信頼区間を求めてグラフを作成する（母分散既知）

ここで作成した「母平均の 99%信頼区間の上限と下限についての折れ線グラフ」は，例題 9-2 で作成した「母平均の 95%信頼区間の上限と下限についての折れ線グラフ」と比べると，区間の幅が大きいことが確認できます．

一般に，ある標本から*母分散*を使って算出される*母平均*の信頼区間の幅は，信頼係数が大きいほど大きくなります．

第 10 章　母平均の区間推定（母分散未知）

10.1　t 分布

例題 10-1　自由度を変えた t 分布のグラフを作成する

自由度 1 の t 分布のグラフ，自由度 3 の t 分布のグラフ，自由度 5 の t 分布のグラフ，および，自由度 100 の t 分布のグラフを，同一グラフエリア内にそれぞれ作成して比べてみましょう．

例題 10-1　自由度を変えた t 分布のグラフを作成する

Point!

t 分布

前章で学習したように，正規分布 $N(\mu, \sigma^2)$ をとる母集団から大きさ n の標本を無作為抽出するとき，母分散 σ^2 はわかっているとすると，次の式から母平均 μ の 95%信頼区間を求めることができます（標本平均を $\overline{X}$ とします）．

$$\overline{X} - \text{「標準正規分布における上側 2.5\%点」} \times \sqrt{\frac{\sigma^2}{n}} \leq \mu$$

$$\leq \overline{X} + \text{「標準正規分布における上側 2.5\%点」} \times \sqrt{\frac{\sigma^2}{n}}$$

例題 10-1　自由度を変えた t 分布のグラフを作成する

Point!

t 分布

しかし，母分散 σ^2 がわからない場合は，これを使うことができません．

そこで，母平均の区間推定をおこなう際に母分散がわからない場合は，母分散 σ^2 の代わりに不偏分散 s^2 を使います．

そして，不偏分散 s^2 を使って母平均の区間推定をおこなう場合には不確実性が増すので，標準正規分布よりもすその厚い t 分布を使います．

例題 10-1　自由度を変えた t 分布のグラフを作成する

Point!

t 分布

t 分布のグラフは，左右対称であり，標準正規分布のグラフと形がよく似ています．

また，t 分布は自由度を決めればただ 1 つに定まり，自由度によってグラフの位置は変わりませんが，グラフの形は変わります．

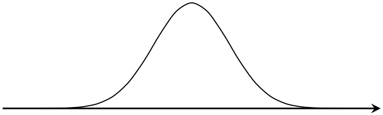
例題 10-1　自由度を変えた t 分布のグラフを作成する

Point!

t 分布

t 分布のグラフは，自由度が大きくなるにつれて，山が高く，すそがうすくなり，標準正規分布のグラフに近づきます．

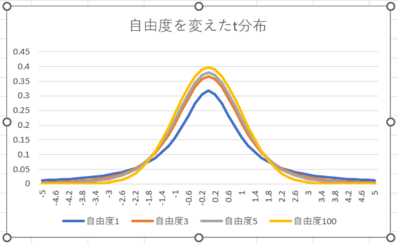
確率変数 X の確率分布が自由度 n の t 分布であるとき，X は自由度 n の t 分布に従うといいます．



自由度 30 の t 分布

例題 10-1　自由度を変えた t 分布のグラフを作成する

X	自由度1	自由度3	自由度5	自由度100	標準正規分布
-5	0.0122427	0.0042194	0.0017574	5.08E-06	
-4.8	0.0132408	0.0048784	0.0021523	1.128E-05	
-4.6	0.0143642	0.0056672	0.0026505	2.456E-05	
-4.4	0.0156341	0.0066163	0.0032826	5.229E-05	
-4.2	0.0170767	0.007765	0.004089	0.0001088	
-4	0.0187241	0.0091634	0.0051237	0.0002212	
-3.8	0.0206159	0.010876	0.0064588	0.0004382	
-3.6	0.0228016	0.0129866	0.0081908	0.0008455	
-3.4	0.0253431	0.0156041	0.0104487	0.0015872	
-3.2	0.0283194	0.0188706	0.0134057	0.0028953	
-3	0.031831	0.022972	0.0172926	0.0051261	
-2.8	0.0360079	0.0281516	0.0224155	0.0087991	
-2.6	0.0410193	0.0347266	0.0291757	0.0146282	
-2.4	0.0470873	0.0431076	0.0380897	0.0235284	
-2.2	0.0545051	0.0538183	0.0498034	0.0365772	
-2	0.063662	0.0675097	0.0650903	0.0549086	



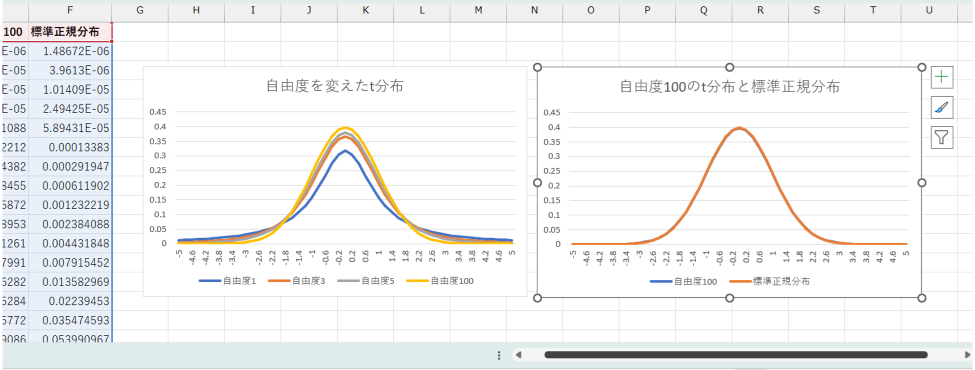
Sheet1

自由度が大きくなるにつれて，山が高く，すそがうすくなることが確認できます．

例題 10-2　t 分布のグラフと標準正規分布のグラフを比較する

自由度 100 の t 分布のグラフ，および，標準正規分布のグラフを，同一グラフエリア内にそれぞれ作成して比べてみましょう．

例題 10-2　t 分布のグラフと標準正規分布のグラフを比較する



2 つのグラフは重なって見え，ほとんどちがいがわからないことが確認できます．

例題 10-3　t 分布の累積分布関数の逆関数の値を求める

確率変数 X が自由度 3 の t 分布に従うとして，X の確率密度関数，累積分布関数の値をそれぞれ求めましょう．また，次をそれぞれ求めましょう．

下位 5%以下に入るのはどの値以下か，

上位 5%以上に入るのはどの値以上か，

下位 2.5%以下に入るのはどの値以下か，

上位 2.5%以上に入るのはどの値以上か，

下位 0.5%以下に入るのはどの値以下か，

上位 0.5%以上に入るのはどの値以上か

例題 10-3　t 分布の累積分布関数の逆関数の値を求める

	A	B	C	D	E	F	G	H	I
1	X	確率密度関数	累積分布関数		自由度3のt分布				
2	-5	0.004219354	0.007696219		下位5%以下に入るのはどの値以下か				-2.35336
3	-4.8	0.00487843	0.008604055		上位5%以上に入るのはどの値以上か				2.353363
4	-4.6	0.005667195	0.009656223		下位2.5%以下に入るのはどの値以下か				-3.18245
5	-4.4	0.006616349	0.010881607		上位2.5%以上に入るのはどの値以上か				3.182446
6	-4.2	0.007765021	0.012316039		下位0.5%以下に入るのはどの値以下か				-5.84091
7	-4	0.009163361	0.014004228		上位0.5%以上に入るのはどの値以上か				5.840909
8	-3.8	0.010875996	0.016002295						
9	-3.6	0.012986623	0.018381104						
10	-3.4	0.015604119	0.021230662						
11	-3.2	0.018870614	0.024665921						
12	-3	0.022972037	0.028834443						
13	-2.8	0.028151623	0.03392645						
14	-2.6	0.034726608	0.040187911						
15	-2.4	0.043107595	0.047937241						
16	-2.2	0.053818288	0.057585976						
17	-2	0.067509661	0.069662984						

Sheet1

自由度 3 の t 分布において，上側 5%点は 2.35，上側 2.5%点は 3.18，上側 0.5%点は 5.84 であることがわかります．

10.2　母平均の信頼区間（母分散未知）

例題 10-4　母平均の 95%信頼区間を求める（母分散未知）

ある試験の点数のデータを母集団とし，そこから無作為に抽出された大きさ 4 の標本により，母平均 μ の 95%信頼区間を求めましょう．ただし，母集団は正規分布をとるとし，母分散 σ^2 はわかっていないとします．

例題 10-4　母平均の 95%信頼区間を求める（母分散未知）

Point!

母平均の 95%信頼区間（母分散未知）

正規分布 $N(\mu, \sigma^2)$ をとる母集団から大きさ n の標本を無作為抽出するとき，母分散 σ^2 はわかっていないとすると，次の式から母平均 μ の 95%信頼区間を求めることができます（標本平均を $\overline{X}$，不偏分散を s^2 とします）．

$$\overline{X} - \text{「自由度}\,(n-1)\,\text{の}\,t\,\text{分布における上側 2.5\%点」} \times \sqrt{\frac{s^2}{n}} \leq \mu$$
$$\leq \overline{X} + \text{「自由度}\,(n-1)\,\text{の}\,t\,\text{分布における上側 2.5\%点」} \times \sqrt{\frac{s^2}{n}}$$

例題 10-4　母平均の 95%信頼区間を求める（母分散未知）

母平均 μ の 95%信頼区間はおおよそ

$$36.3 \leq \mu \leq 90.7$$

となります.

	A	B	C	D	E	F	G
1	母分散 σ^2=?		標本平均	63.5			
2			不偏分散	291.6667			
3	標本						
4	58		自由度3のt分布における上側2.5%点			3.182446305	
5	80						
6	42		母平均 μ の95%信頼区間		36.32469	≤ 母平均 μ ≤	90.67531
7	74						
8							
9							

例題 10-5　標本サイズを大きくして母平均の 95%信頼区間を求める（母分散未知）

ある試験の点数のデータを母集団とし，そこから無作為に抽出された大きさ 100 の標本により，母平均 μ の 95%信頼区間を求めましょう．ただし，母集団は正規分布をとるとし，母分散 σ^2 はわかっていないとします．

そして，例題 10-4 で求めた「大きさ 4 の標本から算出した母平均 μ の 95%信頼区間」と，ここで求めた「大きさ 100 の標本から算出した母平均 μ の 95%信頼区間」を比較してみましょう．

例題 10-5　標本サイズを大きくして母平均の 95%信頼区間を求める（母分散未知）

母平均 μ の 95%信頼区間はおおよそ

$$60.1 \leq \mu \leq 66.9$$

となります．

	A	B	C	D	E	F	G
1	母分散 σ^2=?		標本平均	63.5			
2			不偏分散	291.7071			
3	標本						
4	72		自由度99のt分布における上側2.5%点			1.984216952	
5	96						
6	48		母平均 μ の95%信頼区間		60.11107	≦ 母平均 μ ≦	66.88893
7	52						
8	76						

例題 10-5　標本サイズを大きくして母平均の 95%信頼区間を求める（母分散未知）

ここで求めた「大きさ 100 の標本から算出した母平均 μ の 95%信頼区間」：

$$60.1 \leq \mu \leq 66.9$$

は，例題 10-4 で求めた「大きさ 4 の標本から算出した母平均 μ の 95%信頼区間」：

$$36.3 \leq \mu \leq 90.7$$

に比べて，幅が小さいことが確認できます（標本平均は互いに等しく，不偏分散も互いにほぼ等しいです）．

例題 10-5　標本サイズを大きくして母平均の 95%信頼区間を求める（母分散未知）

前章では，母分散がわかっている場合に，標本サイズが大きいほど，その標本から算出される母平均の信頼区間の幅は小さくなることを確認しました．母分散がわかっていない場合でも，不偏分散が同じであるとすると，同様のことが成り立ちます．

例題 10-6　母平均の 90%信頼区間を求める（母分散未知）

例題 10-4 において，母平均 μ の 95%信頼区間を，90%信頼区間に変更しましょう．

そして，例題 10-4 で求めた「母平均 μ の 95%信頼区間」と，ここで求めた「母平均 μ の 90%信頼区間」を比較してみましょう．

例題 10-6　母平均の 90%信頼区間を求める（母分散未知）

Point!

母平均の 90%信頼区間（母分散未知）

正規分布 $N(\mu, \sigma^2)$ をとる母集団から大きさ n の標本を無作為抽出するとき，母分散 σ^2 はわかっていないとすると，次の式から母平均 μ の 90%信頼区間を求めることができます（標本平均を $\overline{X}$，不偏分散を s^2 とします）．

$$\overline{X} - \text{「自由度}\,(n-1)\,\text{の}\,t\,\text{分布における上側 5\%点」} \times \sqrt{\frac{s^2}{n}} \leq \mu$$

$$\leq \overline{X} + \text{「自由度}\,(n-1)\,\text{の}\,t\,\text{分布における上側 5\%点」} \times \sqrt{\frac{s^2}{n}}$$

例題 10-6　母平均の 90%信頼区間を求める（母分散未知）

母平均 μ の 90%信頼区間はおおよそ

$$43.4 \leq \mu \leq 83.6$$

となります．

	A	B	C	D	E	F	G
1	母分散 σ^2=?		標本平均	63.5			
2			不偏分散	291.6667			
3	標本						
4	58		自由度3のt分布における上側5%点			2.353363435	
5	80						
6	42		母平均 μ の90%信頼区間		43.40433	≤ 母平均 μ ≤	83.59567
7	74						
8							
9							

例題 10-6　母平均の 90%信頼区間を求める（母分散未知）

ここで求めた「母平均 μ の 90%信頼区間」：

$$43.4 \leq \mu \leq 83.6$$

は，例題 10-4 で求めた「母平均 μ の 95%信頼区間」：

$$36.3 \leq \mu \leq 90.7$$

に比べて，幅が小さいことが確認できます．

例題 10-7　母平均の 99%信頼区間を求める（母分散未知）

例題 10-4 において，母平均 μ の 95%信頼区間を，99%信頼区間に変更しましょう．

そして，例題 10-4 で求めた「母平均 μ の 95%信頼区間」と，ここで求めた「母平均 μ の 99%信頼区間」を比較してみましょう．

例題 10-7　母平均の 99%信頼区間を求める（母分散未知）

Point!

母平均の 99%信頼区間（母分散未知）

正規分布 $N(\mu, \sigma^2)$ をとる母集団から大きさ n の標本を無作為抽出するとき，母分散 σ^2 はわかっていないとすると，次の式から母平均 μ の 99%信頼区間を求めることができます（標本平均を $\overline{X}$，不偏分散を s^2 とします）．

$$\overline{X} - \text{「自由度}\ (n-1)\ \text{の}\ t\ \text{分布における上側 0.5\%点」} \times \sqrt{\frac{s^2}{n}} \leq \mu$$
$$\leq \overline{X} + \text{「自由度}\ (n-1)\ \text{の}\ t\ \text{分布における上側 0.5\%点」} \times \sqrt{\frac{s^2}{n}}$$

例題 10-7　母平均の 99%信頼区間を求める（母分散未知）

母平均 μ の 99%信頼区間はおおよそ

$$13.6 \leq \mu \leq 113.4$$

となります．

	A	B	C	D	E	F	G
1	母分散 σ^2=?		標本平均	63.5			
2			不偏分散	291.6667			
3	標本						
4	58		自由度3のt分布における上側0.5%点			5.84090931	
5	80						
6	42		母平均 μ の99%信頼区間		13.62374	≦ 母平均 μ ≦	113.3763
7	74						
8							
9							
10							

例題 10-7　母平均の 99%信頼区間を求める（母分散未知）

ここで求めた「母平均 μ の 99%信頼区間」：

$$13.6 \leq \mu \leq 113.4$$

は，例題 10-4 で求めた「母平均 μ の 95%信頼区間」：

$$36.3 \leq \mu \leq 90.7$$

に比べて，幅が大きいことが確認できます．

例題 10-7　母平均の 99%信頼区間を求める（母分散未知）

前章では，母分散がわかっている場合に，ある標本から母分散を使って算出される母平均の信頼区間の幅は，信頼係数が大きいほど大きくなることを確認しました．母分散がわかっておらず，不偏分散を使って算出する場合でも，同様のことが成り立ちます．

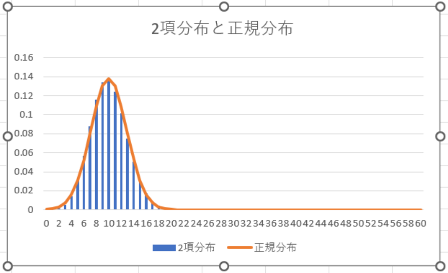
第 11 章　母比率の区間推定

11.1　2 項分布の正規近似

例題 11-1　2 項分布の成功率を変えて正規分布と比較する

まず，例題 5-3 において，2 項分布 $B(60, 1/6)$ のグラフに正規分布 $N(10, 25/3)$ $(= N(10, 8.333\cdots))$ のグラフが近似していることを再確認しましょう．

	A	B	C
1	X	2項分布	正規分布
2	0	1.77E-05	0.000343
3	1	0.000213	0.001071
4	2	0.001256	0.00297
5	3	0.004858	0.007306
6	4	0.013847	0.015938
7	5	0.031016	0.030836
8	6	0.056863	0.052915
9	7	0.087731	0.080534
10	8	0.116244	0.10871
11	9	0.134327	0.13015
12	10	0.137013	0.138198
13	11	0.124557	0.13015
14	12	0.101722	0.10871
15	13	0.075118	0.080534
16	14	0.050436	0.052915
17	15	0.030934	0.030836



例題 11-1　2 項分布の成功率を変えて正規分布と比較する

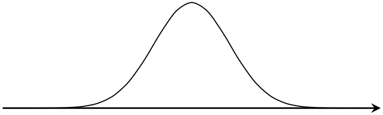
そして，2 項分布の成功率を $1/6$ から $1/100$ に変更し（つまり，$B(60, 1/6)$ を $B(60, 1/100)$ に変更し），それにともない，近似する正規分布も修正します．

成功率が小さくなったときのグラフの変化を見て，それによって近似の精度がどのように変化するか確認してみましょう．

例題 11-1　2 項分布の成功率を変えて正規分布と比較する

2 項分布の正規近似

試行回数 n，成功率 p の 2 項分布 $B(n,p)$ は，試行回数 n が大きくなるにつれて，期待値 np，分散 $np(1-p)$ の正規分布 $N(np, np(1-p))$ に近づきます．



正規分布

例題 11-1　2 項分布の成功率を変えて正規分布と比較する

Point!

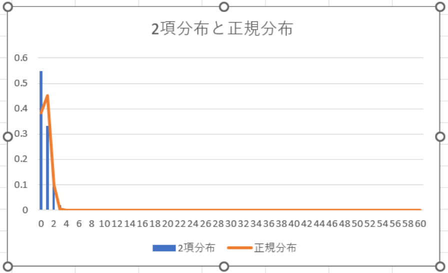
2 項分布の正規近似

ここで，正規近似とは，ある分布を正規分布とみなす近似のことをいいます．2 項分布の正規近似とは，2 項分布を正規分布とみなす近似のことになります．

例題 11-1　2 項分布の成功率を変えて正規分布と比較する

期待値と分散を変更したことにより，C 列が正規分布 $N(3/5, 297/500)$ に修正されます．

	A	B	C
1	X	2項分布	正規分布
2	0	0.547157	0.382307
3	1	0.33161	0.452404
4	2	0.098813	0.099427
5	3	0.019297	0.004058
6	4	0.002778	3.08E-05
7	5	0.000314	4.33E-08
8	6	2.91E-05	1.13E-11
9	7	2.27E-06	5.5E-16
10	8	1.52E-07	4.96E-21
11	9	8.85E-09	8.31E-27
12	10	4.56E-10	2.58E-33
13	11	2.09E-11	1.49E-40
14	12	8.64E-13	1.6E-48
15	13	3.22E-14	3.19E-57
16	14	1.09E-15	1.18E-66
17	15	3.38E-17	8.13E-77



Sheet1

例題 11-1　2 項分布の成功率を変えて正規分布と比較する

グラフが変化し，2 項分布と正規分布のズレがめだつようになりました．

つまり，2 項分布における成功率を 0 に近くすると（0.5 から遠くすると），正規近似の精度が下がりました．

例題 11-1　2 項分布の成功率を変えて正規分布と比較する

このように，2 項分布を正規分布で近似する際，試行回数が同じなら，成功率が 0.5 から遠いほど近似の精度は下がります．また，成功率が 0.5 に近いほど近似の精度は上がります．

例題 11-2　2 項分布の試行回数を変えて正規分布と比較する

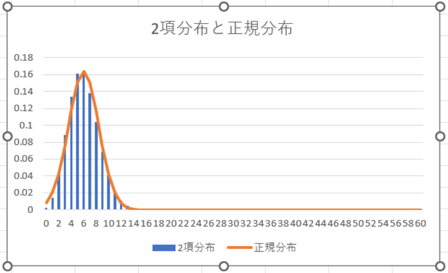
例題 11-1 において，2 項分布の試行回数を 60 から 600 に変更し（つまり，$B(60, 1/100)$ を $B(600, 1/100)$ に変更し），それにともない，近似する正規分布も修正します．

グラフの変化を見て，近似の精度がどのように変化するか確認してみましょう

例題 11-2　2 項分布の試行回数を変えて正規分布と比較する

期待値と分散を変更したことにより，C 列が正規分布 $N(6, 297/50)$ に修正されます．

X	2項分布	正規分布
0	0.002405	0.007906
1	0.014576	0.019957
2	0.044096	0.04257
3	0.088785	0.076737
4	0.13385	0.116893
5	0.161161	0.150474
6	0.161432	0.163688
7	0.13837	0.150474
8	0.103603	0.116893
9	0.068836	0.076737
10	0.041093	0.04257
11	0.022263	0.019957
12	0.011038	0.007906
13	0.005043	0.002647
14	0.002136	0.000749



例題 11-2　2 項分布の試行回数を変えて正規分布と比較する

グラフが変化し，例題 5-3 と同じように，2 項分布と正規分布のズレがめだたなくなりました．

このように，2 項分布を正規分布で近似する際，成功率が同じなら，試行回数が大きいほど近似の精度は上がります．

例題 11-2　2 項分布の試行回数を変えて正規分布と比較する

試行回数 n，成功率 p について，np または $n(1-p)$ の値が小さい場合は，2 項分布と正規分布のズレがめだちます（例題 11-1 では $np = 60 \times 1/100 = 0.6$ であり小さい）．

成功率 p が 0 または 1 に近いときに正規分布とみなすには，試行回数 n の値をかなり大きくしないといけないことに注意しましょう（例題 11-2 では $np = 600 \times 1/100 = 6$ となり大きくなった）．

11.2　母比率の信頼区間

例題 11-3　母比率の 95%信頼区間を求める

N 市在住の大学生 1000 人を対象に，パスポートを保有しているかどうかについてアンケートをとったところ,「保有している」と答えた割合（標本比率）は 30%でした．大学生のパスポート保有率（母比率）の 95%信頼区間を求めましょう.
ここで，標本サイズ 1000 は十分大きく，標本比率が正規分布に従うとします.

例題 11-3　母比率の 95%信頼区間を求める

Point!

母比率，標本比率

母集団において，ある性質をもっている割合は，母比率とよばれます．

母集団がある性質をもっている割合（母比率）を p とすると，母集団から無作為抽出された「大きさ n の標本におけるその性質をもつ個数 X」は，「成功率が p であるベルヌーイ試行を n 回おこなうときに成功する回数 X」と解釈することができ，2 項分布 $B(n,p)$ に従います．

この X を n で割った「X/n」は，標本比率とよばれます．

例題 11-3　母比率の 95%信頼区間を求める

中心極限定理により，標本サイズ n が大きくなるにつれて，標本比率「X/n」の分布は正規分布 $N(p, p(1-p)/n)$ に近づきます（p は母比率）．

例題 11-3　母比率の 95%信頼区間を求める

Point!

母比率の 95%信頼区間

母集団から大きさ n の標本を無作為抽出するとき，次の式から母比率 p の 95%信頼区間を求めることができます（標本比率を $\hat{p}$ とします）．ただし，標本サイズ n が十分大きく，標本比率は正規分布に従うとします．

$$\hat{p} - \text{「標準正規分布における上側 2.5\%点」} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p$$

$$\leq \hat{p} + \text{「標準正規分布における上側 2.5\%点」} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

例題 11-3　母比率の 95%信頼区間を求める

母比率 p の 95%信頼区間はおおよそ

$$27.2\% \leq p \leq 32.8\%$$

となります．

	A	B	C	D	E
1	標本サイズn	1000			
2	標本比率	30%			
3					
4	標準正規分布における上側2.5%点			1.959963985	
5					
6	母比率pの95%信頼区間		27.2%	≤ 母比率p ≤	32.8%
7					
8					
9					

例題 11-4　標本サイズを大きくして母比率の 95%信頼区間を求める

N 市在住の大学生 10000 人を対象に，パスポートを保有しているかどうかについてアンケートをとったところ,「保有している」と答えた割合（標本比率）は 30%でした．N 市在住の大学生のパスポート保有率（母比率）の 95%信頼区間を求めましょう．
ここで，標本サイズ 10000 は十分大きく，標本比率が正規分布に従うとします．

例題 11-4　標本サイズを大きくして母比率の 95%信頼区間を求める

例題 11-3 において，標本サイズを変更しましょう．

そして，例題 11-3 で求めた「大きさ 1000 の標本から算出される母比率 p の 95%信頼区間」と，ここで求めた「大きさ 10000 の標本から算出される母比率 p の 95%信頼区間」を比較してみましょう．

例題 11-4　標本サイズを大きくして母比率の 95%信頼区間を求める

母比率 p の 95%信頼区間はおおよそ

$$29.1\% \leq p \leq 30.9\%$$

となります．

	A	B	C	D	E
1	標本サイズn	10000			
2	標本比率	30%			
3					
4	標準正規分布における上側2.5%点			1.959963985	
5					
6	母比率pの95%信頼区間		29.1%	≤ 母比率p ≤	30.9%
7					
8					
9					

例題 11-4　標本サイズを大きくして母比率の 95%信頼区間を求める

ここで求めた「大きさ 10000 の標本から算出される母比率 p の 95%信頼区間」：

$$29.1\% \leq p \leq 30.9\%$$

は，例題 11-3 で求めた「大きさ 1000 の標本から算出される母比率 p の 95%信頼区間」：

$$27.2\% \leq p \leq 32.8\%$$

に比べて，幅が小さいことが確認できます．

例題 11-4　標本サイズを大きくして母比率の 95%信頼区間を求める

一般に，標本比率が同じなら，標本サイズが大きいほど，その標本から算出される母比率の信頼区間の幅は小さくなります．

例題 11-5　標本比率を変えて母比率の 95%信頼区間を求める

T 市在住の大学生 1000 人を対象に，パスポートを保有しているかどうかについてアンケートをとったところ,「保有している」と答えた割合（標本比率）は 50%でした．T 市在住の大学生のパスポート保有率（母比率）の 95%信頼区間を求めましょう．
ここで，標本サイズ 1000 は十分大きく，標本比率が正規分布に従うとします．

例題 11-5　標本比率を変えて母比率の 95%信頼区間を求める

例題 11-3 において，標本比率を変更しましょう．

そして，例題 11-3 で求めた「標本比率 30%の標本から算出される母比率 p の 95%信頼区間」と，ここで求めた「標本比率 50%の標本から算出される母比率 p の 95%信頼区間」を比較してみましょう．

例題 11-5　標本比率を変えて母比率の 95%信頼区間を求める

母比率 p の 95%信頼区間はおおよそ

$$46.9\% \leq p \leq 53.1\%$$

となります．

	A	B	C	D	E
1	標本サイズn	1000			
2	標本比率	50%			
3					
4	標準正規分布における上側2.5%点			1.959963985	
5					
6	母比率pの95%信頼区間		46.9%	≤ 母比率p ≤	53.1%
7					
8					
9					

例題 11-5　標本比率を変えて母比率の 95%信頼区間を求める

ここで求めた「標本比率 50%の標本から算出される母比率 p の 95%信頼区間」：

$$46.9\% \leq p \leq 53.1\%$$

は，例題 11-3 で求めた「標本比率 30%の標本から算出される母比率 p の 95%信頼区間」：

$$27.2\% \leq p \leq 32.8\%$$

に比べて，幅が大きいことが確認できます．

例題 11-5　標本比率を変えて母比率の 95%信頼区間を求める

一般に，標本サイズが同じなら，標本比率が 0.5 に近いほど，その標本から算出される母比率の信頼区間の幅は大きくなります．

例題 11-6　標本比率を求めて母比率の 95%信頼区間を求める

あるコインを 400 回投げて出た面についてのデータ（標本）にもとづき，このコインで表が出る比率（母比率）の 95%信頼区間を求めましょう.「0」を「裏」,「1」を「表」と解釈することにします.
ここで，標本サイズ 400 は十分大きく，標本比率が正規分布に従うとします.

例題 11-6　標本比率を求めて母比率の 95%信頼区間を求める

母比率 p の 95%信頼区間はおおよそ

$$55.5\% \leq p \leq 65.0\%$$

となります．

	A	B	C	D	E	F	G
1	標本		**標本サイズn**	400			
2	0		**標本比率**	60.3%			
3	1						
4	1		**標準正規分布における上側2.5%点**			1.959963985	
5	0						
6	1		**母比率pの95%信頼区間**		**55.5%**	**≤ 母比率p ≤**	**65.0%**
7	1						
8	0						
9	1						

例題 11-7　母比率の 90%信頼区間を求める

例題 11-6 において，このコインで表が出る比率（母比率）の 95%信頼区間を，90%信頼区間に変更しましょう．

そして，例題 11-6 で求めた「母比率 p の 95%信頼区間」と，ここで求めた「母比率 p の 90%信頼区間」を比較してみましょう．

例題 11-7　母比率の 90%信頼区間を求める

Point!

母比率の 90%信頼区間

母集団から大きさ n の標本を無作為抽出するとき，次の式から母比率 $\boldsymbol{p}$ の 90%信頼区間を求めることができます（標本比率を $\widehat{p}$ とします）．ただし，標本サイズ n が十分大きく，標本比率は正規分布に従うとします．

$$\widehat{p} - \text{「標準正規分布における上側 5\%点」} \times \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} \leq p$$

$$\leq \widehat{p} + \text{「標準正規分布における上側 5\%点」} \times \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}}$$

例題 11-7　母比率の 90%信頼区間を求める

母比率 p の 90%信頼区間はおおよそ

$$56.2\% \leq p \leq 64.3\%$$

となります．

	A	B	C	D	E	F	G
1	標本		標本サイズn	400			
2	0		標本比率	60.3%			
3	1						
4	1		標準正規分布における上側5%点			1.644853627	
5	0						
6	1		母比率pの90%信頼区間		56.2%	≦ 母比率p ≦	64.3%
7	1						
8	0						
9	1						
10	0						
11	1						

例題 11-7　母比率の 90%信頼区間を求める

ここで求めた「母比率 p の 90%信頼区間」：

$$56.2\% \leq p \leq 64.3\%$$

は，例題 11-6 で求めた「母比率 p の 95%信頼区間」：

$$55.5\% \leq p \leq 65.0\%$$

に比べて，幅が小さいことが確認できます．

例題 11-8　母比率の 99%信頼区間を求める

例題 11-6 において，このコインで表が出る比率（母比率）の 95%信頼区間を，99%信頼区間に変更しましょう．

そして，例題 11-6 で求めた「母比率 p の 95%信頼区間」と，ここで求めた「母比率 p の 99%信頼区間」を比較してみましょう．

例題 11-8　母比率の 99%信頼区間を求める

母比率の 99%信頼区間

Point!

母集団から大きさ n の標本を無作為抽出するとき，次の式から母比率 p の 99%信頼区間を求めることができます（標本比率を $\widehat{p}$ とします）．ただし，標本サイズ n が十分大きく，標本比率は正規分布に従うとします．

$$\widehat{p} - \text{「標準正規分布における上側 0.5\%点」} \times \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} \leq p$$

$$\leq \widehat{p} + \text{「標準正規分布における上側 0.5\%点」} \times \sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}}$$

例題 11-8　母比率の 99%信頼区間を求める

母比率 p の 99%信頼区間はおおよそ

$$53.9\% \leq p \leq 66.6\%$$

となります．

	A	B	C	D	E	F	G
1	標本		**標本サイズn**	400			
2	0		**標本比率**	60.3%			
3	1						
4	1		**標準正規分布における上側0.5%点**			2.575829304	
5	0						
6	1		**母比率pの99%信頼区間**		**53.9%**	**≤ 母比率p ≤**	**66.6%**
7	1						
8	0						
9	1						
10	0						
11	1						

例題 11-8　母比率の 99%信頼区間を求める

ここで求めた「母比率 p の 99%信頼区間」：

$$53.9\% \leq p \leq 66.6\%$$

は，例題 11-6 で求めた「母比率 p の 95%信頼区間」：

$$55.5\% \leq p \leq 65.0\%$$

に比べて，幅が大きいことが確認できます．

例題 11-8　母比率の 99%信頼区間を求める

ある標本から算出される母比率の信頼区間の幅は，信頼係数が大きいほど大きくなります．

第 12 章　母比率の検定

12.1　仮説検定の考え方

例題 12-1　成功する回数が近似的に従う正規分布を求める

あるコイン を 400 回投げて出た面についてのデータ（標本）にもとづき，このコインが正しいコインであるかどうかについて，検定をおこないます．

例題 12-1　成功する回数が近似的に従う正規分布を求める

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	標本		**帰無仮説：**	このコインを投げたとき表の出る確率（母比率）は						である（このコインは正しいコインである）						
2	0		**対立仮説：**	このコインを投げたとき表の出る確率（母比率）は						でない（このコインは正しいコインでない）						
3	1															
4	1		**標本サイズn**													
5	0															
6	1		帰無仮説を仮定すると，表の出る回数Xは，			**試行回数**		**，成功率**		**の2項分布**に従う						
7	1		したがって，表の出る回数Xは，		**期待値**		**，分散**		**の正規分布**に従う							
8	0															
9	1		**表が出た回数x**													
10	0		**z**		（検定統計量）											
11	1															
12	1		**標準正規分布における上側2.5%点**													
13	0															
14	1		**有意水準5%の両側検定**					成り立てば帰無仮説を棄却しない →					≤		≤	
15	1															
16	0															
17	0															

Sheet1

成功する回数が従う正規分布を使っておこなう母比率の検定の元データ

例題 12-1　成功する回数が近似的に従う正規分布を求める

そのためまず，このコインが正しいコインであると仮定した際に，400 回投げたときに表の出る回数 X が従う正規分布の期待値と分散を求めましょう．
ここで,「0」を「裏」,「1」を「表」と解釈することにします．また，標本サイズ 400 は十分大きく，X が正規分布に従うとします．

例題 12-1　成功する回数が近似的に従う正規分布を求める

Point! **仮説検定**

仮説検定（または，検定）とは，母集団がとる分布の母数についての仮説を立てたうえで，実際に得られている標本がそこから抽出される確率にもとづき，その仮説が成り立ちそうもないと判断できるかどうかを検証することをいいます．

例題 12-1　成功する回数が近似的に従う正規分布を求める

Point! 仮説検定

その際，最初に立てる仮説のことは帰無仮説とよばれます．そして，それが成り立ちそうもないと判断することを「帰無仮説を棄却する」といいます．

帰無仮説に対する仮説であり，帰無仮説が棄却された場合に採択される仮説のことは対立仮説とよばれます．

例題 12-1　成功する回数が近似的に従う正規分布を求める

この例題の場合，帰無仮説は

　「このコインは正しいコインである」

ということになり，対立仮説は

　「このコインは正しいコインでない」

ということになります．

例題 12-1　成功する回数が近似的に従う正規分布を求める

「このコインは正しいコインである」という帰無仮説のもと，400 回投げたときに表の出る回数 X が従う正規分布を求め，実際に起きた

「表が出た回数が 241 であること」

が，ありえないほど起こりにくいことであるのか，そうでもないのかを判断します．

例題 12-1　成功する回数が近似的に従う正規分布を求める

- ありえないほど起こりにくいことなら，帰無仮説を棄却し，対立仮説を採択することにします．つまり,「このコインは正しいコインでない」といえそうと判断することにします．
- ありえないほど起こりにくいことではないのなら,「帰無仮説を棄却しない」ことにします．この場合は,「このコインは正しいコインでない」とは判断しないことにします．

例題 12-1　成功する回数が近似的に従う正規分布を求める

この例題では，帰無仮説のもとで，このコインを 400 回投げたときに表の出る回数 X が従う正規分布を求めるところまでやってみましょう．

例題 12-1　成功する回数が近似的に従う正規分布を求める

なお，表の出る回数 X の分布は，「0」と「1」からなる母集団から，大きさ 400 の標本の無作為抽出をくり返すときの，その標本に含まれる「1」の個数（= 標本に含まれるデータの合計）の分布と解釈できます．

帰無仮説を仮定すると，$\boldsymbol{X}$ の分布は，試行回数 $\mathbf{400}$，成功率 $\mathbf{1/2}$ の $\mathbf{2}$ 項分布 $\boldsymbol{B}(\mathbf{400}, \mathbf{1/2})$ になります．

例題 12-1　成功する回数が近似的に従う正規分布を求める

ここで，例題 5-1，補足 8-1 で確認したように，試行回数 n，成功率 p の 2 項分布 $B(n, p)$ は，試行回数 n が大きくなるにつれて，期待値 np，分散 $np(1-p)$ の正規分布 $N(np, np(1-p))$ に近づきます（中心極限定理）．

これより，帰無仮説のもとでの X の分布は，近似的に，期待値 $(400 \cdot 1/2 =)$ 200，分散 $(400 \cdot 1/2 \cdot (1-1/2) =)$ 100 の正規分布 $N(200, 100)$ になります．

例題 12-1　成功する回数が近似的に従う正規分布を求める

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	標本		**帰無仮説：**	このコインを投げたとき表の出る確率（母比率）は					0.5	である（このコインは正しいコインである）						
2	0		**対立仮説：**	このコインを投げたとき表の出る確率（母比率）は					0.5	でない（このコインは正しいコインでない）						
3	1															
4	1		**標本サイズn**	400												
5	0															
6	1		帰無仮説を仮定すると，表の出る回数Xは，			**試行回数**	**400**	，**成功率**	**0.5**	**の2項分布**に従う						
7	1		したがって，表の出る回数Xは，		**期待値**	**200**	，**分散**	**100**	**の正規分布**に従う							
8	0															
9	1		**表が出た回数x**													
10	0		**z**		（検定統計量）											
11	1															
12	1		**標準正規分布における上側2.5%点**													
13	0															
14	1		**有意水準5%の両側検定**					成り立てば帰無仮説を棄却しない →					≤		≤	
15	1															
16	0															
17	0															

Sheet1

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

あるコインを 400 回投げて出た面についてのデータ（標本）にもとづき，このコインが正しいコインであるかどうかについて，有意水準 5%で両側検定をおこないます．

例題 12-1 では，このコインが正しいコインであると仮定したうえで，400 回投げたときに表の出る回数 X が従う正規分布の期待値 200 と分散 100 を求めました．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

帰無仮説が正しいと仮定したときの X が従う正規分布の期待値と分散がわかったので，実際に起きた「表が出た回数が 241 であること」が，この帰無仮説のもとではありえないほど起こりにくいことであるのか，または，そうでもないのかを判断しましょう．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

帰無仮説が正しいと仮定したとき，400 回投げたときに表の出る回数 X が従う正規分布の期待値 200 と分散 100 がわかったので，X は次のように標準化することができ，Z は標準正規分布 $N(0, 1^2)$ に従います．

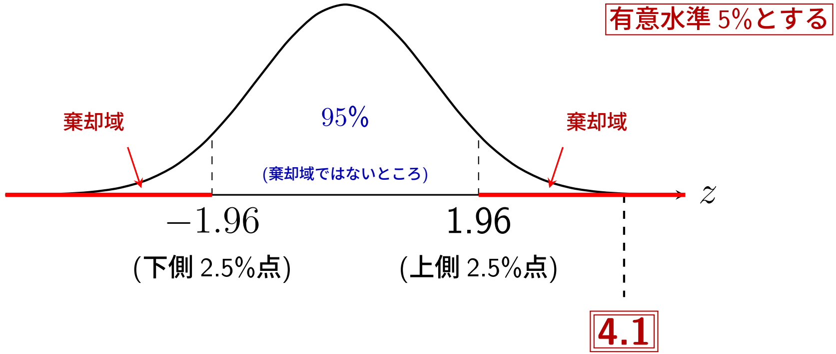
$$\boldsymbol{Z = \frac{X - 200}{\sqrt{100}}} \qquad \left(Z = \frac{X - \text{期待値}}{\text{標準偏差}} \right)$$

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

この Z のような，仮説検定で使用する統計量は検定統計量とよばれます．

そして，実際に表が出た回数 x（$=241$）の z 値（$=4.1$）を求めて，次のような判断をしましょう．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう



標準正規分布 $N(0,1^2)$ の確率密度関数のグラフ（4.1 は上側 2.5%点 1.96 より大きい）

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

- z 値 4.1 が標準正規分布において，上側 2.5%点より大きい，または，下側 2.5%点より小さいなら，実際に起きた「表が出た回数が 241 であること」が，この帰無仮説のもとではありえないほど起こりにくいことであると判断しましょう．

  そして，帰無仮説を棄却し，対立仮説を採択することにしましょう．つまり,「このコインは正しいコインでない」といえそうと判断します．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

- z 値 4.1 が標準正規分布において，上側 2.5%点より大きくもなく，下側 2.5%点より小さくもないのなら，実際に起きた「表が出た回数が 241 であること」が，この帰無仮説のもとではありえないほど起こりにくいことであるとは判断しないことにしましょう．

  そして，帰無仮説を棄却しないことにしましょう．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

ここで,「上側 2.5%点より大きい，または，下側 2.5%点より小さい」というような，帰無仮説が棄却される範囲のことは，棄却域とよばれます.

棄却域が統計検定量の分布の両側にあるような検定は，両側検定といわれます.

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

この例題において，棄却するかどうかを判定する基準として，確率 5%（← 上側 2.5%と下側 2.5%を合わせて）を使いました．

このように,「どれくらい小さな確率をありえないほど起こりにくいとするか」をあらかじめ定めた基準のことを有意水準とよびます．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	標本		**帰無仮説：**	このコインを投げたとき表の出る確率（母比率）は					0.5	である（このコインは正しいコインである）						
2	0		**対立仮説：**	このコインを投げたとき表の出る確率（母比率）は					0.5	でない（このコインは正しいコインでない）						
3	1															
4	1		**標本サイズn**	400												
5	0															
6	1		帰無仮説を仮定すると，表の出る回数Xは，			**試行回数**	**400**	**，成功率**	**0.5**	**の2項分布**に従う						
7	1		したがって，表の出る回数Xは，		**期待値**	**200**	**，分散**	**100**	**の正規分布**に従う							
8	0															
9	1		**表が出た回数x**	241												
10	0		**z**	**4.1**	（検定統計量）											
11	1															
12	1		**標準正規分布における上側2.5%点**			1.959964										
13	0															
14	1		**有意水準5%の両側検定**		棄却する			成り立てば帰無仮説を棄却しない →				-1.95996	≤	**4.1**	≤	1.959964
15	1															
16	0															
17	0															

Sheet1　+

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

帰無仮説を棄却し，対立仮説を採用することにします.

つまり,「このコインは正しいコインでない」といえそうと判断することにします.

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

Point! **例題 12-1 と例題 12-2 の仮説検定の手順のまとめ**

(1) 仮説を立てます．

帰無仮説：このコインを投げたとき表の出る確率（母比率）は $1/2$ である．

対立仮説：このコインを投げたとき表の出る確率（母比率）は $1/2$ でない．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

Point! **例題 12-1 と例題 12-2 の仮説検定の手順のまとめ**

(2) 有意水準を決めます.

5%とします.

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

Point!

例題 12-1 と例題 12-2 の仮説検定の手順のまとめ

(3) 検定統計量を決めます.

- 帰無仮説を仮定すると,「このコインを 400 回投げたときに表の出る回数 X は，試行回数 400，成功率 $1/2$ の 2 項分布 $B(400, 1/2)$ に従う」ということになります.
- 一般に，**試行回数 n，成功率 p の 2 項分布 $B(n, p)$ は，試行回数 n が大きくなるにつれて，期待値 np，分散 $np(1-p)$ の正規分布 $N(np, np(1-p))$ に近づきます.**

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

Point!

例題 12-1 と例題 12-2 の仮説検定の手順のまとめ

- これより，帰無仮説のもとでの X の分布は，近似的に，期待値 200，分散 100 の正規分布 $N(200, 100)$ になります．
- X を標準化した
$$Z = \frac{X - 200}{\sqrt{100}}$$
を検定統計量とします．Z は標準正規分布に従います．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

Point!

例題 12-1 と例題 12-2 の仮説検定の手順のまとめ

(4) 帰無仮説を棄却する条件を決めます．

標準正規分布を使って，両側検定をおこないます．得られている標本から算出される検定統計量（← 次の (5) で求める値）が，標準正規分布における上側 2.5%点より大きい，または，下側 2.5%点より小さいなら，帰無仮説を棄却することにします．

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

Point! 例題 12-1 と例題 12-2 の仮説検定の手順のまとめ

(5) 得られている標本から検定統計量を算出します.

$$z = \frac{241 - 200}{\sqrt{100}} = 4.1$$

例題 12-2　成功する回数が近似的に従う正規分布を使って母比率の検定をおこなう

例題 12-1 と例題 12-2 の仮説検定の手順のまとめ

(6) 帰無仮説を棄却するかしないか結論を出します．

手順 5 で求めた統計量 4.1 は「標準正規分布における上側 2.5%点」より大きいので，帰無仮説を棄却します．

12.2　母比率の検定

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

例題 12-2 と同じ検定を，標本比率が従う正規分布を考えることによりおこなっても，同じ結果になることを確認しましょう．

あるコインを 400 回投げて出た面についてのデータ（標本）にもとづき，このコインが正しいコインであるかどうかについて，有意水準 5%で両側検定をおこないます．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

帰無仮説が正しいと仮定したとき，このコインを 400 回投げたときに表の出る回数 X の分布についてまでは，すでに求められています（例題 12-1 で求めました）．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

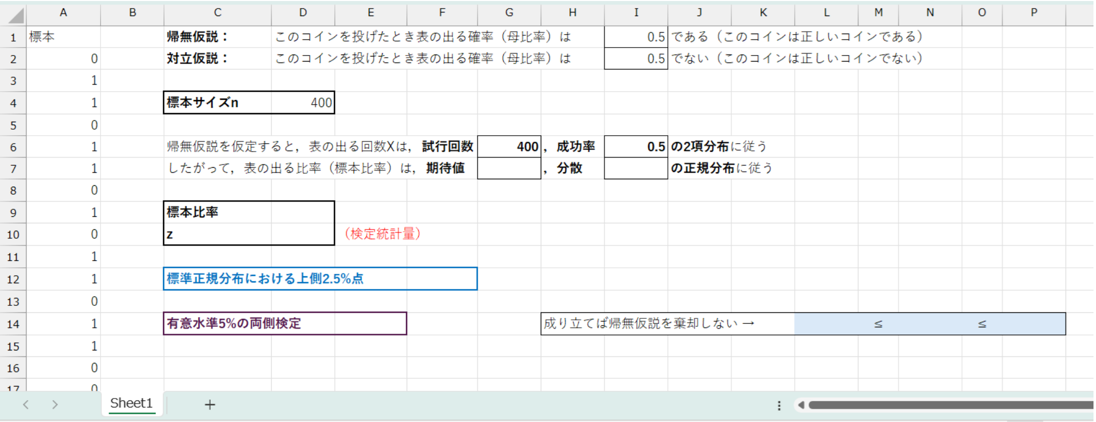
例題 12-2 では，表が出た回数を標準化して検定をしましたが，ここでは，標本比率を標準化して検定をしてみます．

そのためまず，このコインが正しいコインであると仮定した際に，400 回投げたときに表の出る比率（標本比率）が従う正規分布の期待値と分散を求めましょう．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

そして，実際に起きた「標本比率が $241/400$ であること」が，この帰無仮説のもとではありえないほど起こりにくいことであるのか，または，そうでもないのかを判断しましょう．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	標本		**帰無仮説：**	このコインを投げたとき表の出る確率（母比率）は					0.5	である（このコインは正しいコインである）						
2	0		**対立仮説：**	このコインを投げたとき表の出る確率（母比率）は					0.5	でない（このコインは正しいコインでない）						
3	1															
4	1		**標本サイズn**	400												
5	0															
6	1		帰無仮説を仮定すると，表の出る回数Xは，**試行回数**				**400**	，**成功率**	**0.5**	**の2項分布**に従う						
7	1		したがって，表の出る比率（標本比率）は，**期待値**					，**分散**		**の正規分布**に従う						
8	0															
9	1		**標本比率**													
10	0		**z**		（検定統計量）											
11	1															
12	1		**標準正規分布における上側2.5%点**													
13	0															
14	1		**有意水準5%の両側検定**					成り立てば帰無仮説を棄却しない →					≤		≤	
15	1															
16	0															
17	0															

標本比率が従う正規分布を使っておこなう母比率の検定の元データ

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

例題 12-1 において，帰無仮説を仮定すると，このコインを 400 回投げたときに表の出る回数 X の分布は，試行回数 400，成功率 $1/2$ の 2 項分布 $B(400, 1/2)$ になることを確認しました．

ここで，例題 11-3 で確認したように，X が 2 項分布 $B(n, p)$ に従うとき，試行回数（標本サイズ）n が大きくなるにつれて，標本比率 X/n の分布は正規分布 $B(p, p(1-p)/n)$ に近づきます（中心極限定理）．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

これより，帰無仮説のもとでの標本比率の分布は，近似的に，期待値 $1/2$，分散 $(1/2 \cdot (1-1/2)/400 =)$ $1/1600$ の正規分布 $N(1/2, 1/1600)$ になります．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

標本比率 $X/n\ (= X/400)$ を標準化した

$$Z = \frac{X/400 - 1/2}{\sqrt{1/1600}}$$

$$\left(\text{変形すると，例題 12-2 で求めた } Z = \frac{X - 200}{\sqrt{100}} \text{ になる}\right)$$

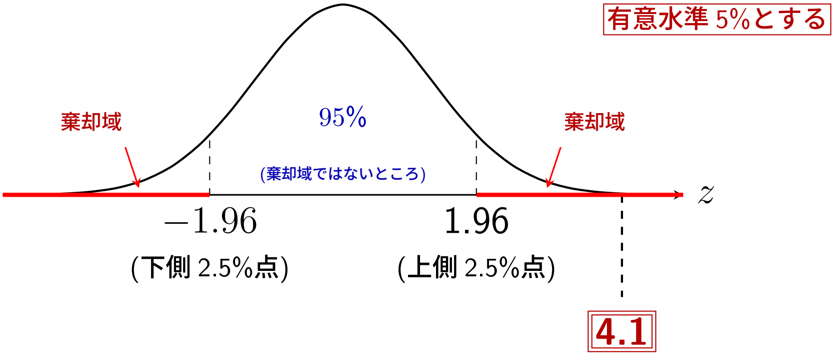
を検定統計量とします．Z は標準正規分布に従います．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	標本		**帰無仮説：**	このコインを投げたとき表の出る確率（母比率）は					0.5	である（このコインは正しいコインである）						
2	0		**対立仮説：**	このコインを投げたとき表の出る確率（母比率）は					0.5	でない（このコインは正しいコインでない）						
3	1															
4	1		**標本サイズn**	400												
5	0															
6	1		帰無仮説を仮定すると，表の出る回数Xは，			**試行回数**	**400**	**，成功率**	**0.5**	**の2項分布**に従う						
7	1		したがって，表の出る比率（標本比率）は，			**期待値**	**0.5**	**，分散**	**0.000625**	**の正規分布**に従う						
8	0															
9	1		**標本比率**	0.6025												
10	0		**z**	**4.1**	（検定統計量）											
11	1															
12	1		**標準正規分布における上側2.5%点**			1.959964										
13	0															
14	1		**有意水準5%の両側検定**		棄却する			成り立てば帰無仮説を棄却しない →				-1.95996	≤	**4.1**	≤	1.959964
15	1															
16	0															
17	0															

Sheet1

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（4.1 は上側 2.5%点 1.96 より大きい）

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

帰無仮説を棄却し，対立仮説を採用することにします.

つまり,「このコインは正しいコインでない」といえそうと判断することにします.

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

標本比率が従う正規分布を用いて検定をおこなっても，得られている標本から算出される検定統計量は，例題 12-2 と同じ値（4.1）になりました．

つまり，例題 12-3 の検定は，表の出る回数が従う正規分布を用いておこなった例題 12-2 の検定と同じ結果になることが確認されました．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

Point!

例題 12-3 の仮説検定の手順のまとめ

(1) 仮説を立てます．

帰無仮説：このコインを投げたとき表の出る確率（母比率）は $1/2$ である．

対立仮説：このコインを投げたとき表の出る確率（母比率）は $1/2$ でない．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

Point! **例題 12-3 の仮説検定の手順のまとめ**

(2) 有意水準を決めます.

5%とします.

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

Point!

例題 12-3 の仮説検定の手順のまとめ

(3) 検定統計量を決めます．

- 帰無仮説を仮定すると,「このコインを 400 回投げたときに表の出る回数 X は，試行回数 400，成功率 $1/2$ の 2 項分布 $B(400, 1/2)$ に従う」ということになります．
- 一般に，$\boldsymbol{X}$ が 2 項分布 $\boldsymbol{B(n,p)}$ に従うとき，試行回数（標本サイズ）$\boldsymbol{n}$ が大きくなるにつれて，標本比率 $\boldsymbol{X/n}$ の分布は正規分布 $\boldsymbol{B(p, p(1-p)/n)}$ に近づきます．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

例題 12-3 の仮説検定の手順のまとめ

- これより，帰無仮説のもとでの標本比率の分布は，近似的に，期待値 $1/2$，分散 $1/1600$ の正規分布 $N(1/2, 1/1600)$ になります．
- 標本比率 X/n $(= X/400)$ を標準化した

$$Z = \frac{X/400 - 1/2}{\sqrt{1/1600}}$$

を検定統計量とします．Z は標準正規分布に従います．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

Point!

例題 12-3 の仮説検定の手順のまとめ

(4) 帰無仮説を棄却する条件を決めます．

標準正規分布を使って，両側検定をおこないます．得られている標本から算出される検定統計量（← 次の (5) で求める値）が，標準正規分布における上側 2.5%点より大きい，または，下側 2.5%点より小さいなら，帰無仮説を棄却することにします．

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

Point!

例題 12-3 の仮説検定の手順のまとめ

(5) 得られている標本から検定統計量を算出します．

$$z = \frac{0.6025 - 1/2}{\sqrt{1/1600}} = 4.1$$

例題 12-3　標本比率が近似的に従う正規分布を使って母比率の検定をおこなう

例題 12-3 の仮説検定の手順のまとめ

(6) 帰無仮説を棄却するかしないか結論を出します.

手順 5 で求めた統計量 4.1 は「標準正規分布における上側 2.5%点」より大きいので，帰無仮説を棄却します.

例題 12-4　母比率の検定を有意水準 5%でおこなう

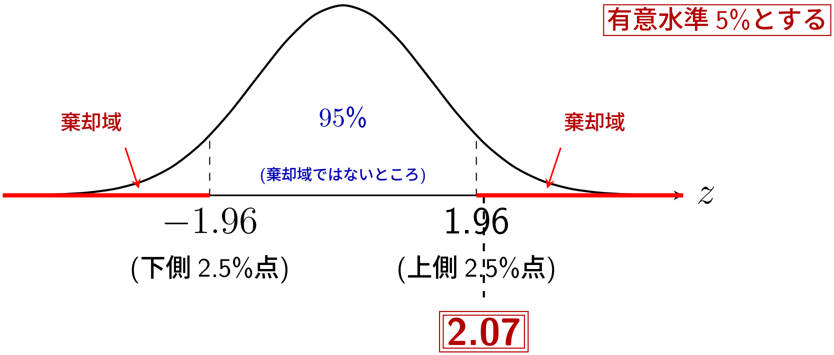
N 大学の学生 1000 人を対象に，パスポートを保有しているかどうかについてアンケートをとったところ,「保有している」と答えた割合（標本比率）は 33%でした．このことにもとづき，大学生のパスポート保有率（母比率）が 30%であるかどうかについて，有意水準 5%で両側検定をおこないます.
ここで，標本サイズ 1000 は十分大きく，標本比率が正規分布に従うとします.

例題 12-4　母比率の検定を有意水準 5%でおこなう

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**帰無仮説：**	大学生のパスポート保有率（母比率）は				0.3	である						
2	**対立仮説：**	大学生のパスポート保有率（母比率）は				0.3	でない						
3													
4	**標本サイズn**	1000											
5													
6	帰無仮説を仮定すると，保有人数Xは，			**試行回数**	**1000**	**，成功率**	**0.3**	**の2項分布**に従う					
7	よって，保有率（標本比率）は，		**期待値**	**0.3**	**，分散**	**0.00021**	**の正規分布**に従う						
8													
9	**標本比率**	0.33											
10	**z**	**2.070197**	（検定統計量）										
11													
12	**標準正規分布における上側2.5%点**			1.959964									
13													
14	**有意水準5%の両側検定**		棄却する		成り立てば帰無仮説を棄却しない →				-1.95996	≤	**2.070197**	≤	1.959964
15													
16													

Sheet1

例題 12-4　母比率の検定を有意水準 5%でおこなう



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（**2.07** は上側 2.5%点 1.96 より大きい）

例題 12-4　母比率の検定を有意水準 5%でおこなう

帰無仮説を棄却し，対立仮説を採用することにします．

つまり,「大学生のパスポート保有率（母比率）は 0.3 でない」といえそうと判断することにします．

例題 12-5　母比率の検定を有意水準 1%でおこなう

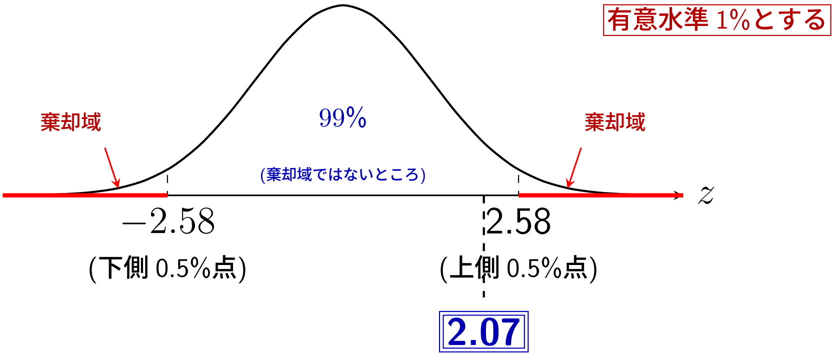
N 大学の学生 1000 人を対象に，パスポートを保有しているかどうかについてアンケートをとったところ,「保有している」と答えた割合（標本比率）は 33%でした．このことにもとづき，大学生のパスポート保有率（母比率）が 30%であるかどうかについて，有意水準 1%で両側検定をおこないます．
ここで，標本サイズ 1000 は十分大きく，標本比率が正規分布に従うとします．

例題 12-5　母比率の検定を有意水準 1%でおこなう

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**帰無仮説：**	大学生のパスポート保有率（母比率）は				0.3	である						
2	**対立仮説：**	大学生のパスポート保有率（母比率）は				0.3	でない						
3													
4	**標本サイズn**	1000											
5													
6	帰無仮説を仮定すると，保有人数Xは，			**試行回数**	**1000**	**，成功率**	**0.3**	**の2項分布**に従う					
7	よって，保有率（標本比率）は，		**期待値**	**0.3**	**，分散**	**0.00021**	**の正規分布**に従う						
8													
9	**標本比率**	0.33											
10	**z**	**2.070197**	（検定統計量）										
11													
12	**標準正規分布における上側0.5%点**			2.575829									
13													
14	**有意水準1%の両側検定**		棄却しない		成り立てば帰無仮説を棄却しない →				-2.57583	≤	**2.070197**	≤	2.575829
15													
16													

Sheet1

例題 12-5　母比率の検定を有意水準 1%でおこなう



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（**2.07** は下側 0.5%点 -2.58 より大きく上側 0.5%点 2.58 より小さい）

例題 12-5　母比率の検定を有意水準 1%でおこなう

帰無仮説を棄却しないことにします.

つまり,「大学生のパスポート保有率（母比率）が 30%でない」とは判断しないことにします.

第 13 章　母平均の検定

13.1　母平均の検定（母分散既知）

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

あるテスト について，A クラス 70 名の生徒の平均点は 47.6 でした．A クラスの生徒の点数はふつうであるといえるかどうかについて，有意水準 5%で両側検定をおこないます．
なお全体では，このテストの点数は期待値 50，分散 100 の正規分布に従っているとします．

例題 13-1　母平均の検定を有意水準 5% でおこなう（母分散既知）

帰無仮説は

「A クラスの点数はふつうである」

ということになり，対立仮説は

「A クラスの点数はふつうでない」

ということになります．

例題 13-1　母平均の検定を有意水準 5% でおこなう（母分散既知）

ここで，例題 8-6 で確認したように，母集団が正規分布 $N(\mu, \sigma^2)$ をとる場合は，標本サイズ n によらず，標本平均の分布は正規分布 $N(\mu, \sigma^2/n)$ になります（正規分布の再生性）．

これより，帰無仮説のもとでの標本平均 $\overline{X}$ の分布は，期待値 50，分散 $(100/70 =)$ $10/7$ の正規分布 $N(50, 10/7)$ になります．

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

標本平均 $\overline{X}$ を標準化した

$$Z = \frac{\overline{X} - 50}{\sqrt{10/7}} \qquad \left(Z = \frac{\overline{X} - \text{期待値}}{\text{標準偏差}} \right)$$

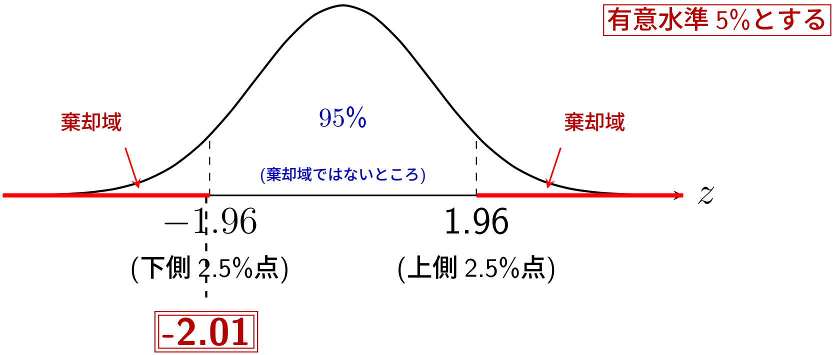
を検定統計量とします． Z は標準正規分布に従います．

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**帰無仮説：**	このテストの平均点（母平均）は			50	である							
2	**対立仮説：**	このテストの平均点（母平均）は			50	でない							
3													
4	**母分散**	100											
5													
6	**標本サイズn**	70											
7													
8	帰無仮説を仮定すると，標本平均は，**期待値**				**50**	**，分散**	**1.428571**	**の正規分布**に従う					
9													
10	**標本平均**	47.6											
11	**z**	**-2.00798**	（検定統計量）										
12													
13	**標準正規分布における上側2.5%点**			1.959964									
14													
15	**有意水準5%の両側検定**		棄却する		成り立てば帰無仮説を棄却しない →				-1.95996	≤	**-2.00798**	≤	1.959964
16													

Sheet1

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（$\mathbf{-2.01}$ は下側 2.5%点 -1.96 より小さい）

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

帰無仮説を棄却し，対立仮説を採用することにします．

つまり,「このテストの平均点（母平均）は 50 でない」（「A クラスの生徒の点数はふつうでない」）といえそうと判断することにします．

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

Point!

例題 13-1 の仮説検定の手順のまとめ

(1) 仮説を立てます.

帰無仮説：このテストの平均点（母平均）は 50 である.

対立仮説：このテストの平均点（母平均）は 50 でない.

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

Point! **例題 13-1 の仮説検定の手順のまとめ**

(2) 有意水準を決めます.

5%とします.

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

Point! 例題 13-1 の仮説検定の手順のまとめ

(3) 検定統計量を決めます．

- 一般に，母集団が正規分布 $N(\mu, \sigma^2)$ をとる場合は，標本サイズ n によらず，標本平均の分布は正規分布 $N(\mu, \sigma^2/n)$ になります．
- これより，帰無仮説のもとでの標本平均 $\overline{X}$ の分布は，期待値 50，分散 $10/7$ の正規分布 $N(50, 10/7)$ になります．

例題 13-1　母平均の検定を有意水準 5% でおこなう（母分散既知）

Point! **例題 13-1 の仮説検定の手順のまとめ**

- 標本平均 $\overline{X}$ を標準化した

$$Z = \frac{\overline{X} - 50}{\sqrt{10/7}}$$

を検定統計量とします．Z は標準正規分布に従います．

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

例題 13-1 の仮説検定の手順のまとめ

(4) 帰無仮説を棄却する条件を決めます．

標準正規分布を使って，両側検定をおこないます．得られている標本から算出される検定統計量（← 次の (5) で求める値）が，標準正規分布における上側 2.5%点より大きい，または，下側 2.5%点より小さいなら，帰無仮説を棄却することにします．

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

Point! **例題 13-1 の仮説検定の手順のまとめ**

(5) 得られている標本から検定統計量を算出します.

$$z = \frac{47.6 - 50}{\sqrt{10/7}} = -2.01 \qquad (-2.01\text{ は概数})$$

例題 13-1　母平均の検定を有意水準 5%でおこなう（母分散既知）

例題 13-1 の仮説検定の手順のまとめ

(6) 帰無仮説を棄却するかしないか結論を出します．

手順 5 で求めた統計量 -2.01 は「標準正規分布における下側 2.5% 点」より小さいので，帰無仮説を棄却します．

例題 13-2　母平均の検定を有意水準 1%でおこなう（母分散既知）

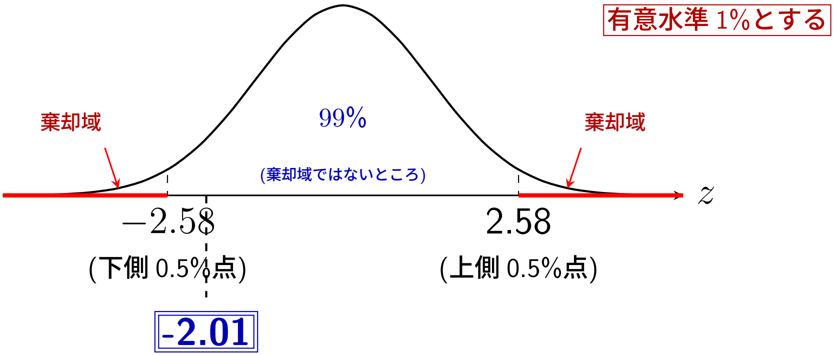
あるテストについて，A クラス 70 名の生徒の平均点は 47.6 でした．A クラスの生徒の点数はふつうであるといえるかどうかについて，有意水準 1%で両側検定をおこないます.
なお全体では，このテストの点数は期待値 50，分散 100 の正規分布に従っているとします.

例題 13-2　母平均の検定を有意水準 1%でおこなう（母分散既知）

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**帰無仮説：**	このテストの平均点（母平均）は			50	である							
2	**対立仮説：**	このテストの平均点（母平均）は			50	でない							
3													
4	**母分散**	100											
5													
6	**標本サイズn**	70											
7													
8	帰無仮説を仮定すると，標本平均は，**期待値**				**50**	**，分散**	**1.428571**	**の正規分布**に従う					
9													
10	**標本平均**	47.6											
11	**z**	**-2.00798**	（検定統計量）										
12													
13	**標準正規分布における上側0.5%点**			2.5758293									
14													
15	**有意水準1%の両側検定**		棄却しない		成り立てば帰無仮説を棄却しない →				-2.57583	≤	**-2.00798**	≤	2.575829
16													

Sheet1

例題 13-2　母平均の検定を有意水準 1%でおこなう（母分散既知）



標準正規分布 $N(0, 1^2)$ の確率密度関数のグラフ（$\mathbf{-2.01}$ は下側 0.5%点 -2.58 より大きく上側 0.5%点 2.58 より小さい）

例題 13-2　母平均の検定を有意水準 1%でおこなう（母分散既知）

帰無仮説を棄却しないことにします.

つまり,「このテストの平均点（母平均）は 50 でない」（「A クラスの生徒の点数はふつうでない」）とは判断しないことにします.

13.2　母平均の検定（母分散未知）

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

ある小学校の 4 年生の D クラス 40 名の児童について，身長の平均値は 136 cm，不偏分散は 25 cm^2 でした．D クラスの児童の身長はふつうであるといえるかどうかについて，有意水準 5%で両側検定をおこないます．
なお全体では，小学 4 年生の身長の分布は正規分布であるとし，平均身長（母平均）は 134 cm であるとします．

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

帰無仮説は

　「D クラスの児童の身長はふつうである」

ということになり，対立仮説は

　「D クラスの児童の身長はふつうでない」

ということになります．

例題 13-3　母平均の検定を有意水準 5% でおこなう（母分散未知）

ここで，補足 10-1 で確認したように，正規分布 $N(\mu, \sigma^2)$ をとる母集団から大きさ n の標本を無作為抽出するとき，次の統計量

$$T = \frac{\overline{X} - \mu}{\sqrt{s^2/n}}$$

は，自由度 $(n-1)$ の t 分布に従います（母分散 σ^2 はわかっていないとし，標本平均を $\overline{X}$，不偏分散を s^2 とします）．

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

これより，帰無仮説のもとで，標本平均 $\overline{X}$ を変換した

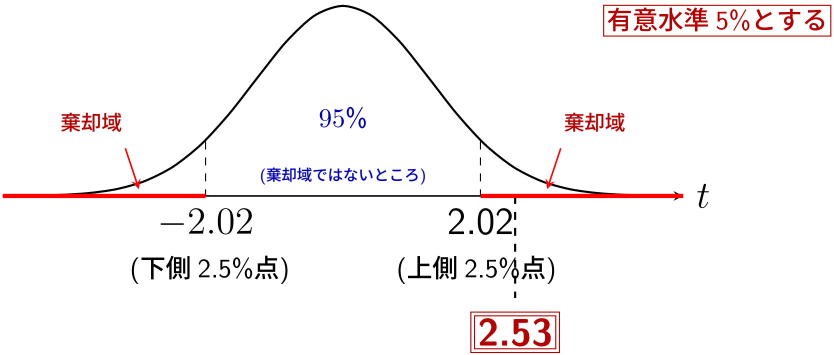
$$T = \frac{\overline{X} - 134}{\sqrt{25/40}}$$

は，自由度 39 の t 分布に従います．この T を検定統計量とします．

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**帰無仮説：**	小学4年生の身長の平均値（母平均）は				134	cmである						
2	**対立仮説：**	小学4年生の身長の平均値（母平均）は				134	cmでない						
3													
4	**標本サイズn**	40											
5													
6	**不偏分散**	25											
7													
8	帰無仮説を仮定すると，標本平均を T へ変換したものは，					**自由度**	**39**	**のt分布**に従う					
9													
10	**標本平均**	136											
11	**t**	**2.529822**	（検定統計量）										
12													
13	**自由度**	**39**	**のt分布における上側2.5%点**			2.022691							
14													
15	**有意水準5%の両側検定**		棄却する		成り立てば帰無仮説を棄却しない →				-2.02269	≤	**2.529822**	≤	2.022691
16													

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）



自由度 39 の t 分布の確率密度関数のグラフ（**2.53** は上側 2.5%点 2.02 より大きい）

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

帰無仮説を棄却し，対立仮説を採用することにします．

つまり,「小学 4 年生の身長の平均値（母平均）は 134 cm でない」（「D クラスの児童の身長はふつうでない」）といえそうと判断することにします．

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

Point! **例題 13-3 の仮説検定の手順のまとめ**

(1) 仮説を立てます．

帰無仮説：小学 4 年生の身長の平均値（母平均）は 134 cm である．

対立仮説：小学 4 年生の身長の平均値（母平均）は 134 cm でない．

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

Point! **例題 13-3 の仮説検定の手順のまとめ**

(2) 有意水準を決めます.

5%とします.

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

Point! 例題 13-3 の仮説検定の手順のまとめ

(3) 検定統計量を決めます．

- 一般に，正規分布 $N(\mu, \sigma^2)$ をとる母集団から大きさ n の標本を無作為抽出するとき，次の統計量

$$T = \frac{\overline{X} - \mu}{\sqrt{s^2/n}}$$

は，自由度 $(n-1)$ の t 分布に従います（母分散 σ^2 はわかっていないとし，標本平均を $\overline{X}$，不偏分散を s^2 とします）．

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

Point! 例題 13-3 の仮説検定の手順のまとめ

- これより，帰無仮説のもとで，標本平均 $\overline{X}$ を変換した
$$T = \frac{\overline{X} - 134}{\sqrt{25/40}}$$
を検定統計量とします．T は，自由度 39 の t 分布に従います．

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

例題 13-3 の仮説検定の手順のまとめ

(4) 帰無仮説を棄却する条件を決めます．

自由度 39 の t 分布を使って，両側検定をおこないます．得られている標本から算出される検定統計量（← 次の (5) で求める値）が，自由度 39 の t 分布における上側 2.5%点より大きい，または，下側 2.5%点より小さいなら，帰無仮説を棄却することにします．

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

Point! 例題 13-3 の仮説検定の手順のまとめ

(5) 得られている標本から検定統計量を算出します.

$$t = \frac{136 - 134}{\sqrt{25/40}} = 2.53 \qquad (2.53 \text{ は概数})$$

例題 13-3　母平均の検定を有意水準 5%でおこなう（母分散未知）

Point!

例題 13-3 の仮説検定の手順のまとめ

(6) 帰無仮説を棄却するかしないか結論を出します．

手順 5 で求めた統計量 2.53 は「自由度 39 の t 分布における上側 2.5% 点」より大きいので，帰無仮説を棄却します．

例題 13-4　母平均の検定を有意水準 1%でおこなう（母分散未知）

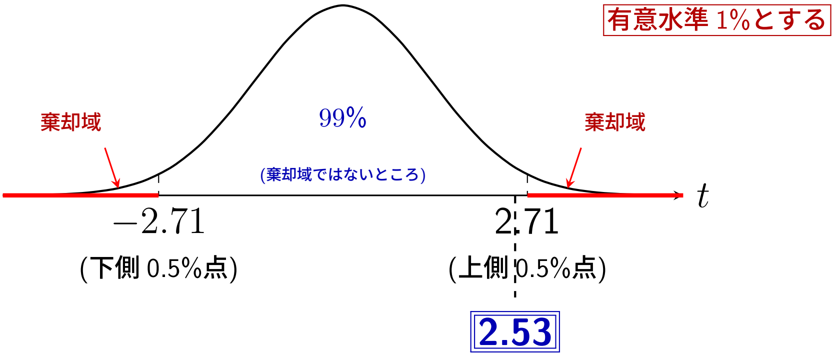
ある小学校の 4 年生の D クラス 40 名の児童について，身長の平均値は 136 cm，不偏分散は 25 cm^2 でした．D クラスの児童の身長はふつうであるといえるかどうかについて，有意水準 1%で両側検定をおこないます．
なお全体では，小学 4 年生の身長の分布は正規分布であるとし，平均身長（母平均）は 134 cm であるとします．

例題 13-4　母平均の検定を有意水準 1%でおこなう（母分散未知）

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**帰無仮説：**	小学4年生の身長の平均値（母平均）は				134	cmである						
2	**対立仮説：**	小学4年生の身長の平均値（母平均）は				134	cmでない						
3													
4	**標本サイズn**	40											
5													
6	**不偏分散**	25											
7													
8	帰無仮説を仮定すると，標本平均は，			**自由度**	**39**	**のt分布**に従う							
9													
10	**標本平均**	136											
11	**t**	**2.529822**	（検定統計量）										
12													
13	**自由度 39**		**のt分布における上側0.5%点**			2.707913							
14													
15	**有意水準1%の両側検定**		棄却しない		成り立てば帰無仮説を棄却しない →				-2.70791	≤	**2.529822**	≤	2.707913
16													

Sheet1

例題 13-4　母平均の検定を有意水準 1%でおこなう（母分散未知）



自由度 39 の t 分布の確率密度関数のグラフ（**2.53** は下側 0.5%点 -2.71 より大きく上側 0.5%点 2.71 より小さい）

例題 13-4　母平均の検定を有意水準 1%でおこなう（母分散未知）

帰無仮説を棄却しないことにします．

つまり，「小学 4 年生の身長の平均値（母平均）は 134 cm でない」（「D クラスの児童の身長はふつうでない」）とは判断しないことにします．