

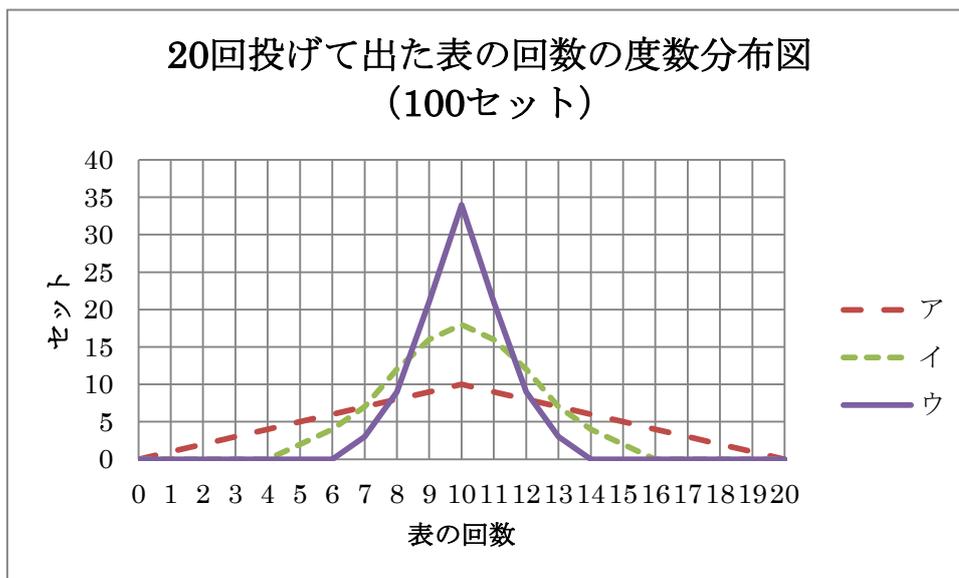
1 はじめに

- ・ 演習結果の紙を持ち帰りたい場合、提出用以外にもう1枚印刷して下さい。
- ・ 演習結果のデータを持ち帰りたい場合は、ログインして各自のフォルダーに保存するか、USBメモリーを持参して保存して下さい。

ビデオ 二項分布と正規分布、正規分布、二項分布シミュレーション (4分)

2 二項分布

演習1 「コインを20回投げる表が出る回数を数える」という試行を100セット行うときの度数の分布はア、イ、ウのどれであると予想しますか？



- ・ 分布

- ある現象が大小さまざまな大きさで起こること。

- ・ 二項分布（にこうぶんぷ）

- ある試行である事象が起こる確率が一定の場合において、その試行みを一定回数行った場合に、その事象が起きる回数の分布。

- 【参考】 コインを 20 回投げたときに表が出る回数の分布

- ◇ 20 回の投げて r 回表が出る確率は ${}_{20}C_r \times 0.5^{20}$ です。

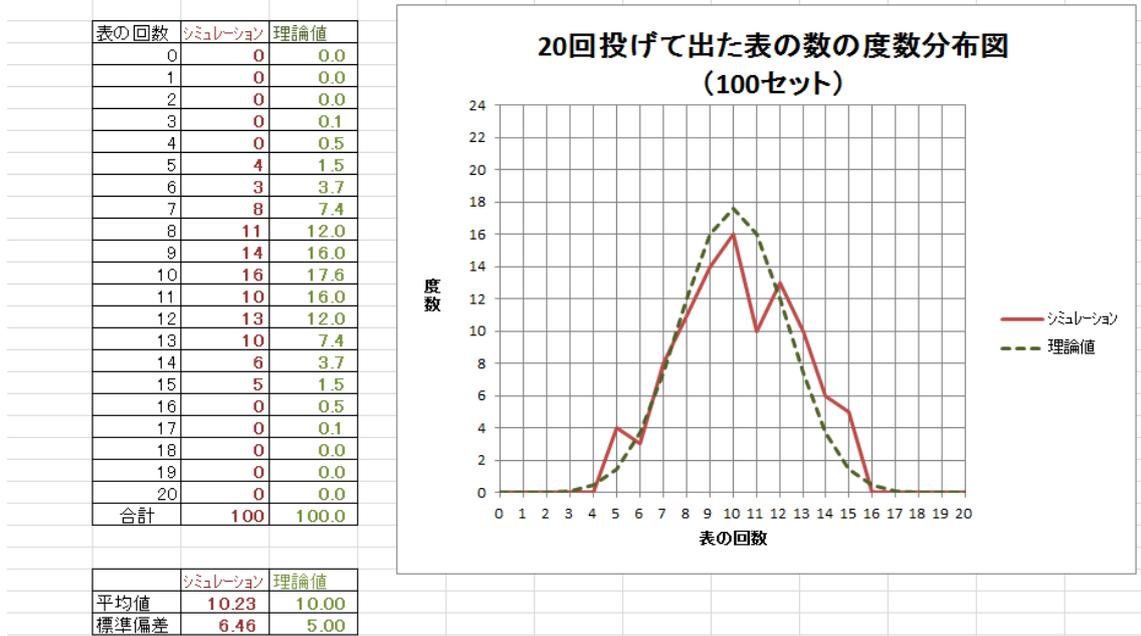
(ここで ${}_{20}C_r = \frac{20!}{(20-r)!r!}$ です。)

演習 2 「コインを 20 回投げる表が出る回数を数える」という試行を Excel で乱数を発生させて 100 セット行って下さい。結果を、度数分布表、度数分布図にするとともに、平均と標準偏差を求めてください。

(ヒント)

- ① このような試みをシミュレーションといいます。
- ② 乱数を発生させます。Excel 関数は =rand() です。
- ③ 乱数の大きさが 0.5 以上だったら 1(表)、0.5 未満なら 0(裏)とみなします。
 - Excel 関数は =if(乱数>=0.5, 1, 0) です。
 - 乱数>=0.5 なら 1、そうでないなら 0 を返します。
- ④ 20 回コインを投げて「表の出た回数」を数えます。
- ⑤ それを 100 回行って、「表の出た回数」に関する度数分布表と度数分布図を作ります。
(注) 下図では、見やすさを考え、棒グラフではなく、折れ線グラフを用います。
- ⑥ 「表の出た回数」の平均値と標準偏差を求めます。

20回投げて出た表の数の度数分布表(100セット)

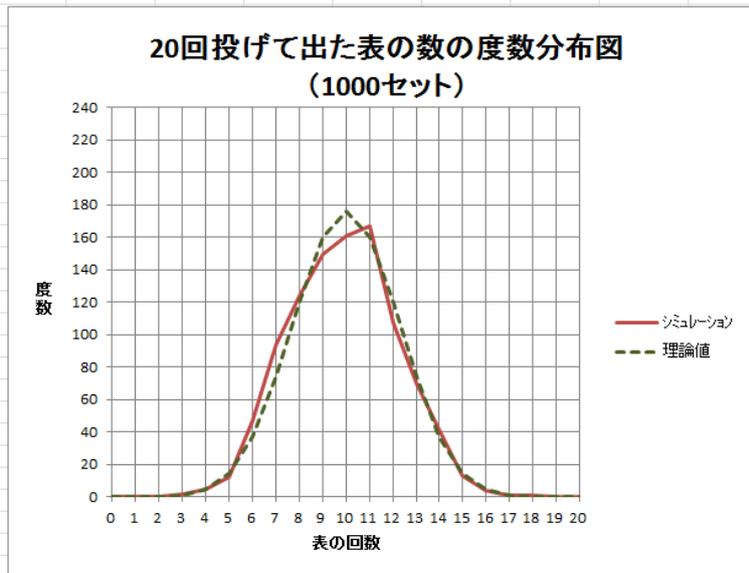


3 大数の法則 (たいすうの法則)

- ・ 試す回数を増やせば増やすほど、その事象の起こる確率は一定の値に近づくという法則。
 - 「20回のコイン投げを100セット」の度数分布 → 理論値から離れている。
 - 「20回のコイン投げを1000セット」の度数分布 → より理論値に近い。

20回投げて出た表の数の度数分布表(1000セット)

表の回数	シミュレーション	理論値
0	0	0.0
1	0	0.0
2	0	0.2
3	2	1.1
4	5	4.6
5	12	14.8
6	47	37.0
7	94	73.9
8	124	120.1
9	150	160.2
10	161	176.2
11	167	160.2
12	107	120.1
13	70	73.9
14	42	37.0
15	13	14.8
16	4	4.6
17	1	1.1
18	1	0.2
19	0	0.0
20	0	0.0
合計	1000	1000.0



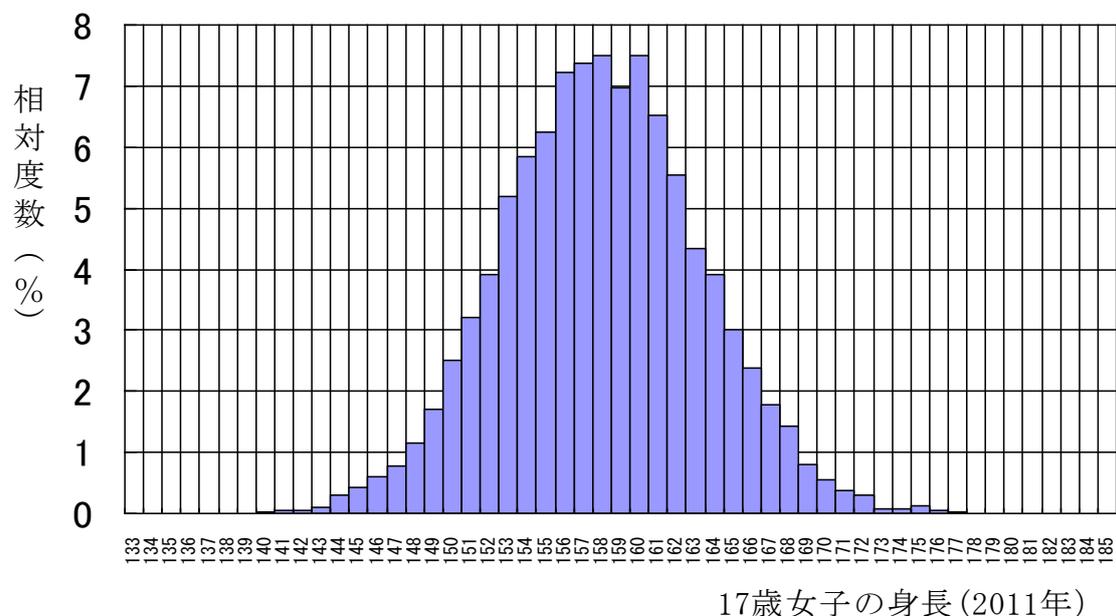
	シミュレーション	理論値
平均値	9.89	10.00
標準偏差	5.29	5.00

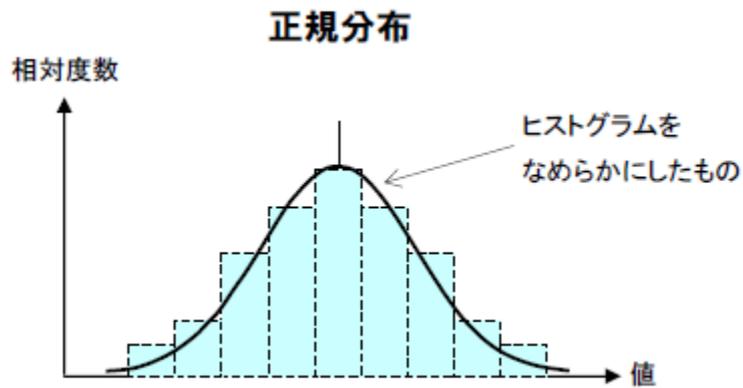
4 正規分布

- ・ 自然現象と社会現象を通じて広くみられる分布の一つです。同じ条件下で得られたバラつきのある数値の発生頻度を調べてみると、中央の値の発生頻度が最も多く、それを中心に対して左右対称で発生頻度が少なくなります。
- ・ 度数分布図は、平均値を頂点とした左右対称の釣鐘型になっています。
 - 例：日本人 17 歳女性の身長分布
- ・ 正規分布はなぜこのような形になるのでしょうか？
 - パチンコを例に考えてみます。パチンコ玉が上から落ちてきて、1つ目の釘に当たる際、右に落ちるか左に落ちるかは、 $1/2$ と考えられます。次の釘にパチンコ玉が当たったら、今度も右・左それぞれ $1/2$ の確率で落ちると考えられます。各場所に落ちたパチンコ玉の数を図にすると釣鐘状に近くなります。
 - このような分布を二項分布といいます。釘の段を増やし、パチンコ玉の数も増やすと、二項分布は正規分布に近づくことが知られています。

参考 文部科学省学校保健調査（2011 年）の結果によると、17 歳女子の身長の相対度数分布図(2011 年)は下図のとおりで、平均値は 158cm、標準偏差は 5.3cm です。

17歳女子の身長の相対度数分布図(2011年)





(注) 度数分布図のことを英語で「ヒストグラム」と言います。

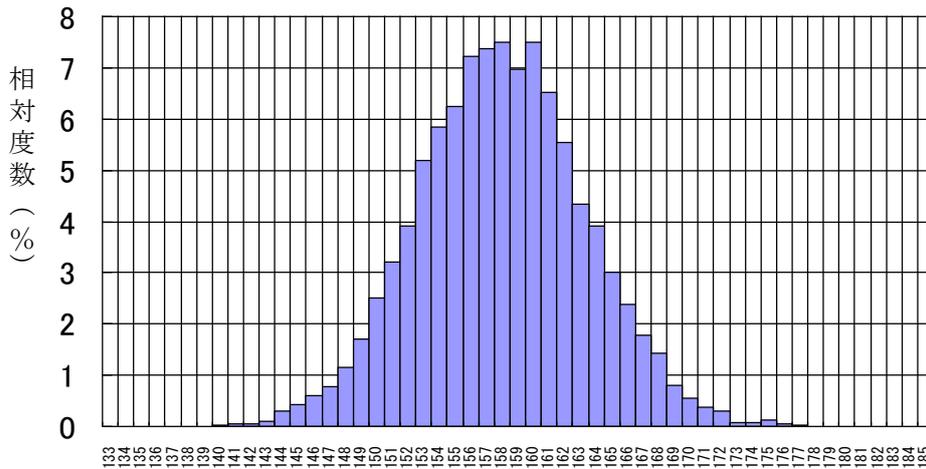
- ・ 自然界で起こる現象の分布には正規分布に当てはまるものが多いです。
 - (例) 睡眠時間、摂取カロリー量、大学入試センター試験の得点など
- ・ 二項分布と正規分布は、似た分布です。
- ・ コイン投げの場合、1セットのコイン投げの回数(現在:20回)を増やし、セット数(現在:100セット)を増やすと正規分布となります。

分布名	例	説明
二項分布	コイン投げ	偶然に左右される
正規分布	身長	遺伝子、食事、運動、睡眠等のたくさんの偶然の要素に左右される。
	肉まんの重さ	原材料や工程での微妙な誤差が重なりあって決まってくる。
	体重	遺伝子、食事、運動、睡眠等のたくさんの偶然に左右される。しかし、偶然の要素が身長ほどには大きくない。

身長と体重はどちらがより正規分布に近いか?

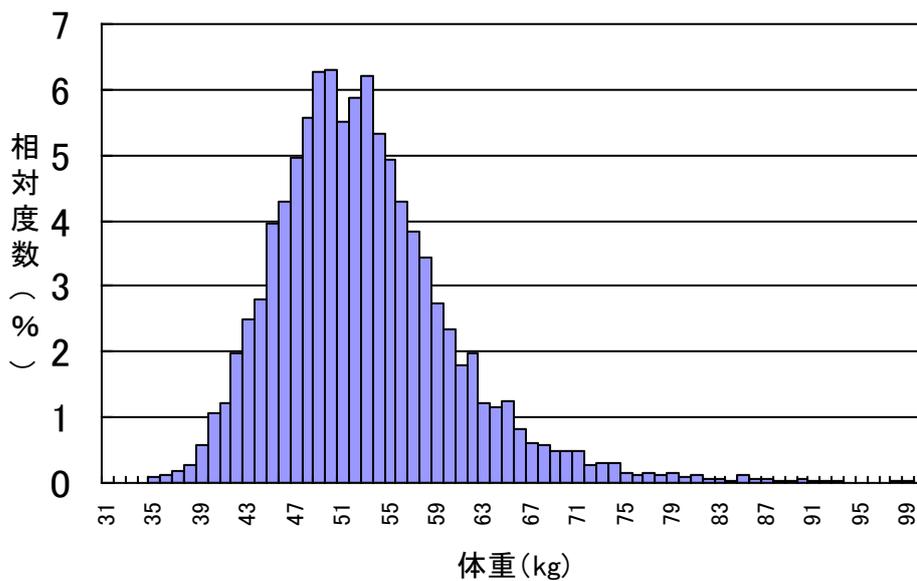
- ・ 身長の分布： 遺伝、生育環境などにより決まり、成長期を過ぎればほぼ変化がない。
- ・ 体重の分布： 遺伝、生育環境のほか、特に、成長期以降では、食生活や運動量などにも左右される。
- ・ すなわち、前者の方が、偶然の結果が作り上げる分布である正規分布により近い。

17歳女子の身長の相対度数分布図(2011年)



17歳女子の身長 (2011年)

17歳女子の体重の相対度数分布図
(文部科学省学校保健調査2011)



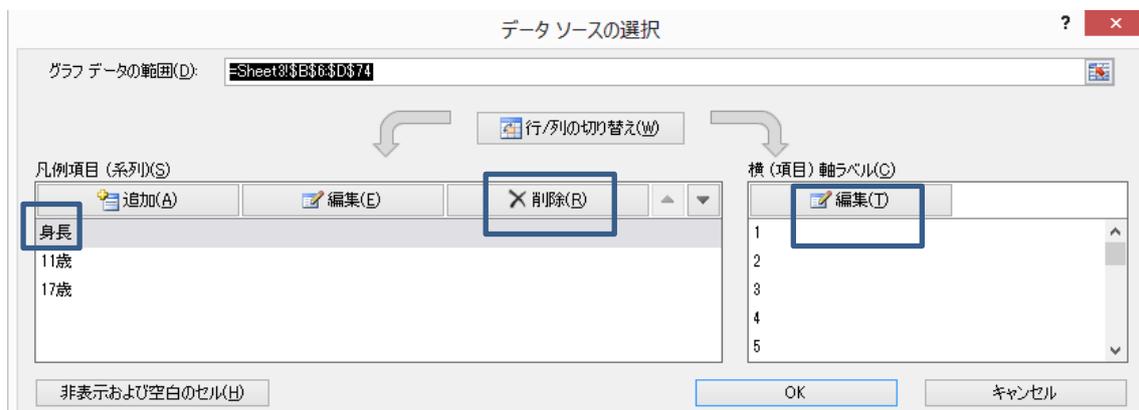
演習 3

- ・ 文部科学省の学校保健統計調査(2013年度)によると、11歳女子と17歳女子の身長
の相対度数分布表は、Excelの表のとおりです。
- ・ 分布の様子はどのように異なるでしょうか？
- ・ 以下の(1)～(3)の作業を進め、提出して下さい。なお、ヘッダーに学籍番号、氏名を
記入することを忘れないで下さい。

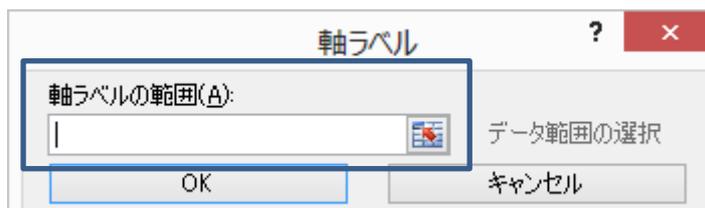
(1) 度数分布図を、折れ線グラフで作成して下さい

(手順)

- ① Excelで「見出し」部分を含める形でデータを選択
- ② 「挿入」タブ→「グラフ」→「折れ線」→「2D折れ線」の一番左のアイコンを選択
→ (グラフができる)
- ③ 「グラフツール」タブ→「デザイン」タブ→「データの選択」
→ (以下の「データソースの選択」ボックスが出てくる)



- ④ 「凡例項目」の「身長」を選択し「削除」をクリック
- ⑤ 「横 (項目) 軸ラベル」の「編集」をクリック
→ (以下の「軸ラベル」のボックスが出てくる)



- ⑥ Excel表で「身長」見出しの「117～184」の数字を選択すると、
自動的に「軸ラベルの範囲」にセルの範囲が代入される→「OK」を押す。

⑦ グラフタイトルを入れる

- 「グラフツール」タブ→「レイアウト」タブ→「グラフタイトル」
- グラフタイトル=「11歳女子と17歳女子の身長分布」

⑧ グラフの縦軸ラベルを入れる

- 「グラフツール」タブ→「レイアウト」タブ→「軸ラベル」→「主縦軸ラベル」
- 縦軸のラベル名=「相対度数 (%)」

⑨ グラフの横軸ラベルを入れる

- 「グラフツール」タブ→「レイアウト」タブ→「軸ラベル」→「主横軸ラベル」
- 横軸のラベル名=「身長(cm)」

(2) 次の空欄を求めて下さい

	平均値①	標準偏差②	①-②	①+②
11歳女子	146.8cm	6.64cm	140.16cm	153.44cm
17歳女子	158.0cm	5.39cm	152.61cm	163.39cm

11歳女子の141cm～153cmの人の全体に占める割合(%)	
17歳女子の153cm～163cmの人の全体に占める割合(%)	

※正規分布の場合この数値の理論値は68.3%となります↑

(3) 11歳女子と17歳女子の身長分布について気づいたことは何ですか？

Excelの所定の欄に簡単に結構ですので書いて下さい。