数科学探究Ｉ「統計学の基礎①」

２年　　組　　号 氏名

これから数時間は「統計学の基礎」を学習します。
統計学はこれから上級学校（大学・短大・専門学校）に進学してから大活躍します。
様々な研究成果を，説得力のあるものにするためには統計学は必須アイテムです。将来のためと思って頑張りましょう。

ただし，今日は数学Ｉの「データの分析」と，中学数学の「標本調査」の復習です (^^;

【問題①】

次の資料は，あるクラスの生徒 10 人の，1 年間に読んだ本の冊数を調べたものです。

18，12，16，19，17，186，21，18，15，14

(1)　平均値を求めなさい。

(2)　中央値を求めなさい。

(3)　このデータの特徴を表す数値（代表値）として，平均値と中央値ではどちらがよいといえるか。理由をつけて答えよ。

【問題②】

箱に白色のビーズがたくさん入っています。手元にある赤色のビーズ 100 個を箱に入れてよくかき混ぜたあと，無作為に取り出すと，75 個のビーズがあり，その中に赤色のビーズが 3 個入っていました。箱の中の白色のビーズの個数を推測しなさい。

【問題③】

池に魚が何匹いるかを調べるのに，次のようにした。この池からかたよりのないように 100 匹の魚を捕えて，それらに印をつけて池にもどす。その後しばらくしてから，もう一度かたよりのないように 300 匹の魚を捕えて調べたら，そのうち 21 匹に印がついていた。この池に魚が何匹いるかを 100 匹単位で推測しなさい。

【問題④】

次のデータは，5 人の生徒の数学のテストの得点 x である。

50，70，90，80，60

(1)　このデータの平均値 $\overline{x}$ を求めよ。

(2)　このデータの分散，標準偏差を求めよ。

(3)　学期成績をつけるために，各テストの得点 x に対して，$y=0.9x+10$ という計算を行った。このとき，新たな値 y について平均，分散，標準偏差を求めよ。

【問題⑤】

あるクラスの10人について，漢字の「書き取り」のテストを行った結果，得点の平均値は64点，標準偏差は7点であった。また，他の2人にも，同じテストを行ったところ，その2人の得点はともに64点であった。12人全員の得点の標準偏差の値は，2人を加える前の値と比較して，大きくなるか，小さくなるか，それとも変わらないかを答えよ。

【問題⑥】

ある高校3年生1クラスの生徒40人について，ハンドボール投げの飛距離のデータを取った。次の図1は，このクラスで最初に取ったデータのヒストグラムである。

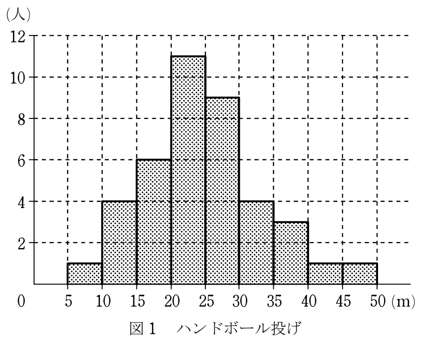


図1　ハンドボール投げ

(1)　次の $\boxed{\text{ア}}$ に当てはまるものを，下の⓪〜⑧のうちから一つ選べ。

この40人のデータの第3四分位数が含まれる階級は，$\boxed{\text{ア}}$ である。

⓪　5 m以上10 m未満　　①　10 m以上15 m未満

②　15 m以上20 m未満　　③　20 m以上25 m未満

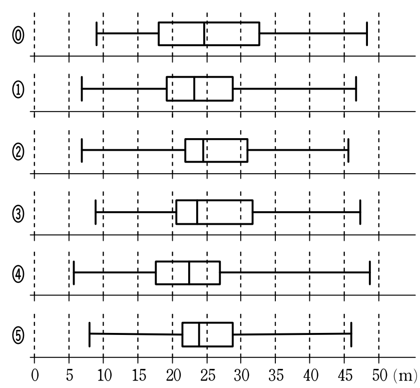
④　25 m以上30 m未満　　⑤　30 m以上35 m未満

⑥　35 m以上40 m未満　　⑦　40 m以上45 m未満

⑧　45 m以上50 m未満

(2)　次の $\boxed{\text{イ}}$ 〜 $\boxed{\text{オ}}$ に当てはまるものを，下の⓪〜⑤のうちから一つずつ選べ。ただし，$\boxed{\text{イ}}$ 〜 $\boxed{\text{オ}}$ の解答の順序は問わない。

このデータを箱ひげ図にまとめたとき，図1のヒストグラムと矛盾するものは，$\boxed{\text{イ}}$，$\boxed{\text{ウ}}$，$\boxed{\text{エ}}$，$\boxed{\text{オ}}$ である。



(3)　次の文章中の $\boxed{\text{カ}}$，$\boxed{\text{キ}}$ に入れるものとして最も適当なものを，下の⓪〜③のうちから一つずつ選べ。ただし，$\boxed{\text{カ}}$，$\boxed{\text{キ}}$ の解答の順序は問わない。

後日，このクラスでハンドボール投げの記録を取り直した。次に示したA〜Dは，最初に取った記録から今回の記録への変化の分析結果を記述したものである。a〜dの各々が今回取り直したデータの箱ひげ図となる場合に，⓪〜③の組合せのうち分析結果と箱ひげ図が矛盾するものは，$\boxed{\text{カ}}$，$\boxed{\text{キ}}$ である。

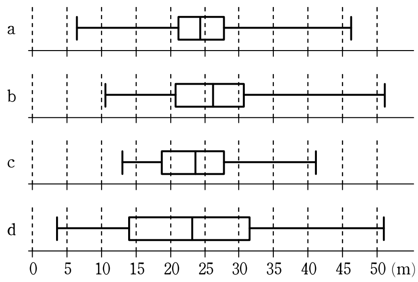
⓪　A−a　　①　B−b　　②　C−c　　③　D−d

A：どの生徒の記録も下がった。

B：どの生徒の記録も伸びた。

C：最初に取ったデータで上位 $\frac{1}{3}$ に入るすべての生徒の記録が伸びた。

D：最初に取ったデータで上位 $\frac{1}{3}$ に入るすべての生徒の記録は伸び，下位 $\frac{1}{3}$ に入るすべての生徒の記録は下がった。



数科学探究Ⅰ「統計学の基礎②」

２年　　組　　号 氏名

「統計学の基礎」第２回です。
　統計学はこれから上級学校（大学・短大・専門学校）に進学してから大活躍します。
　様々な研究成果を，説得力のあるものにするためには統計学は必須アイテムです。将来のためと思って頑張りましょう。

＜平均（期待値）の復習＞

100本のくじがあり，その賞金と本数は下の表のようになっている。

	A賞	B賞	C賞	
賞金(円)	300	100	50	合計
本数(本)	10	30	60	
確率				

このくじを１本引くときの期待できる賞金額（＝期待値or平均）を求めよ。

＜講義メモ　分散＞

下のような確率分布

X	x_1	x_2	………	x_n	計
確率	p_1	p_2	………	p_n	1

において，その平均（期待値）を m とする。

このとき，X の取る値の平均 m との差の２乗の平均を「分散」という。そのルートをとったものが「標準偏差」です。

【問題①】
１個のサイコロを投げるとき，出る目Xの平均E(X)を求めよ。

X	1	2	3	4	5	6	計
確率							

【問題②】
１枚の硬貨を投げるとき，表の出る枚数の平均を求めよ。

【問題③】
２枚の硬貨を同時に投げるとき，表の出る枚数の平均を求めよ。

【例】
１個のサイコロを投げるとき，出る目Xの分散V(X)を求めよ。

【問題④】
１枚の硬貨を投げるとき，表の出る枚数の分散を求めよ。

【問題⑤】
２枚の硬貨を同時に投げるとき，表の出る枚数の分散を求めよ。

平均と分散についての公式

$$E(X+Y)=E(X)+E(Y)$$
$$V(X+Y)=V(X)+V(Y)$$ ※ X と Y が独立のとき

＜例＞

大小２個のサイコロを投げるとき，出る目の和の平均・分散・標準偏差を求めよ。

【問題⑥】

大中小３個のサイコロを投げるとき，出る目の和の平均・分散・標準偏差を求めよ。

【問題⑦】

コインを１０回投げるとき，表の出る枚数の平均・分散・標準偏差を求めよ。

＜終わった人向けおまけ問題＞

1 辺の長さが 1 の正六角形 ABCDEF の頂点から異なる 3 点を選び，これらを頂点とする三角形を作る。

(1)　作られる三角形が正三角形となる確率を求めよ。

(2)　作られる三角形の面積の期待値を求めよ。

答 (1) $\dfrac{1}{10}$　(2) $\dfrac{\sqrt{3}}{4}\cdot\dfrac{6}{20}+\dfrac{\sqrt{3}}{2}\cdot\dfrac{12}{20}+\dfrac{3\sqrt{3}}{4}\cdot\dfrac{2}{20}=\dfrac{9\sqrt{3}}{20}$

表題　　　　　　　　　　　　　　　　　　　　　　　　　　（　　）組（　　）番　名前（　　　　）

＜問題＞
アラビア語を全く知らない人が，このアラビア語単語テスト５０問をテキトーに解くとき，次の各問いに答えよ。

(1)　全問正解する確率を求めよ。（式だけでいいです）

(2)　ちょうど３０問正解する確率を求めよ。（式だけでいいです）

(3)　このテストをたくさんの人が受験したとき，その平均点は何点になると思われますか。

このような，「起こる vs 起こらない」を何回も行う。
「あたる vs あたらない」を何回も行う。

ときの確率の分布のことを $\boxed{}$ といいます。

試行回数が n 回で，
１回の試行で，事象Aの起こる確率が p（起こらない確率 $1-p$）の

＜例＞　１枚のコインを２回投げるときの確率分布

表の出る回数	0	1	2	計
確率				

数科学探究Ⅰ「統計学の基礎③」

２年　　組　　号 氏名

「統計学の基礎」第３回です。
統計学はこれから上級学校（大学・短大・専門学校）に進学してから大活躍します。
様々な研究成果を，説得力のあるものにするためには統計学は必須アイテムです。将来のためと思って頑張りましょう。

＜これまでの復習＞

下のような確率分布

X	x_1	x_2	………	x_n	計
確率	p_1	p_2	………	p_n	1

において，

■平均（期待値）$m=E(X)$
$=x_1p_1+x_2p_2+x_3p_3+\cdots\cdots\cdots+x_np_n$
※「得点×確率」の和

■分散　$V(X)=(x_1-m)^2p_1+(x_2-m)^2p_2+\cdots\cdots+(x_n-m)^2p_n$
※(得点－平均)^{2}の平均

■標準偏差　$\sigma(X)=\sqrt{V(X)}$
※ 標準偏差＝$\sqrt{\text{分散}}$

□平均と分散についての公式

$$E(X+Y)=E(X)+E(Y)$$
$$V(X+Y)=V(X)+V(Y)$$
※ X と Y が独立のとき

本日のテーマ「二項分布」

１個のサイコロを３回投げるとき，３の倍数が出る回数を X とする。
３の倍数が２回でる確率は

３の倍数が r 回でる確率は

ちなみにこの確率変数 X の確率分布は

X(回数)	0	1	2	3	計
P(確率)	${}_3C_0\left(\frac{2}{3}\right)^3$	${}_3C_1\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^2$	${}_3C_2\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^1$	${}_3C_3\left(\frac{1}{3}\right)^3$	1

このような，　「起こる」確率 p
「起こらない」確率 $q=1-p$　のとき，
n 回の試行のうち起こる回数を X としたときの確率分布のことを

といい，　　と書く

＜例＞　１枚の硬貨を５回投げるとき，表の出る回数を X とする。

X は二項分布　　に従う。

たとえば，$P(X=2)=$

【問題①】
１個のサイコロを４回投げるとき，３の倍数の目が出る回数を X とする。X はどんな分布に従うか。また，$P(X=3)$ を求めよ。

【問題②】
１枚の硬貨を１０回投げるとき，表が出る回数を X とする。
(1)　X はどんな分布に従うか。また，$P(X\geqq 9)$ を求めよ。

(2)　このコインを本当に１０回投げたら，表が９回も出た。
このコインは偏りのない普通のコインと言えるかどうか，あなたの意見を理由とともに述べよ。

二項分布の平均・分散・標準偏差

あるコインはちょっと重心がずれており，
表の出る確率は p，裏が出る確率は $q=1-p$ である。

■このコインを１回投げるときの，表がでる回数 X の平均・分散は

X(回数)	0	1	計
P(確率)	q	p	1

$(p+q=1)$

□平均

□分散

■このコインを３回投げるときの，表がでる回数 X について
この X は二項分布 $B(3,\ p)$ に従う。
表の公式　$E(X+Y)=E(X)+E(Y)$，$V(X+Y)=V(X)+V(Y)$　より

□平均

□分散

まとめ

確率変数 X が，二項分布 $B(n, p)$ に従うとき

$E(X)=$　　　　$V(X)=$　　　　$\sigma(X)=$

＜例＞
○，× で答える 10 個の問題に，でたらめに ○，× をつけるとき，そのうちの正解数を X とする。X の期待値，分散と標準偏差を求めよ。

【問題③】
5 つの解答群の中から正しいものを 1 つ選ぶ五者択一式の問題が 20 題ある。でたらめにこの解答をするとき，そのうちの正解数を X とする。X の期待値と分散および標準偏差を求めよ。

※ここまでできれば本日のノルマ達成

＜以下おまけ問題＞

1　さいころを 30 回投げるとき，1 の目の出る回数 X の期待値，分散，標準偏差を求めよ。

解答　期待値 5，分散 $\frac{25}{6}$，標準偏差 $\frac{5\sqrt{6}}{6}$

2　針が上を向く確率が $\frac{3}{5}$ である画びょうを 9 回投げるとき，針が上を向く回数 X の期待値，分散，標準偏差を求めよ。

解答　期待値 $\frac{27}{5}$，分散 $\frac{54}{25}$，標準偏差 $\frac{3\sqrt{6}}{5}$

3　2 % の割合で不良品を含むネジの山から 150 個取り出したとき，それに含まれる不良品の個数を X とする。X の期待値，分散と標準偏差を求めよ。

解答　期待値 3，分散 $\frac{147}{50}$，標準偏差 $\frac{7\sqrt{6}}{10}$

「統計学の基礎」第４回です。
たくさんの試行を行うときの確率を，グラフを用いて求めることができるようになります。

＜これまでの復習＞

□下のような確率分布

X	x_1	x_2	………	x_n	計
確率	p_1	p_2	………	p_n	1

において，

■平均（期待値） $m=E(X)$
$=x_1p_1+x_2p_2+x_3p_3+\cdots\cdots\cdots+x_np_n$
※「得点×確率」の和

■分散　$V(X)=(x_1-m)^2p_1+(x_2-m)^2p_2+\cdots\cdots+(x_n-m)^2p_n$
※(得点－平均)^{2}の平均

■標準偏差　$\sigma(X)=\sqrt{V(X)}$
※ 標準偏差＝$\sqrt{分散}$

□平均と分散についての公式

$$E(X+Y)=E(X)+E(Y)$$
$$V(X+Y)=V(X)+V(Y)$$
※ X と Y が独立のとき

□二項分布
「起こる」確率 p，「起こらない」確率 $q=1-p$　のとき，
n 回の試行のうち起こる回数を X としたときの確率分布のことを二項分布といい，$B(n,\ p)$ と表す。

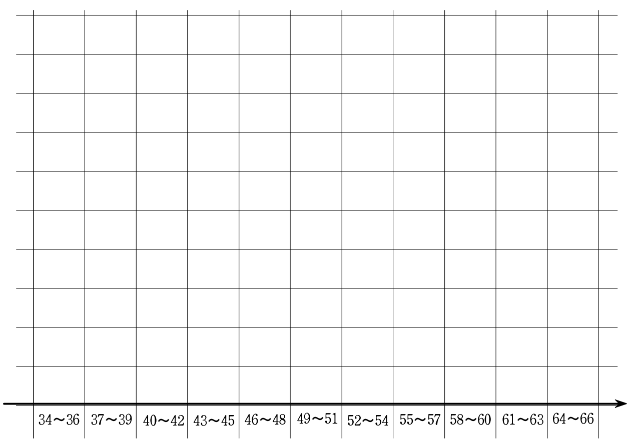
例：サイコロを 10 回なげて，4 の目が出る回数 X は
二項分布 $B\left(10,\ \dfrac{1}{6}\right)$ に従う。

確率変数 X が，二項分布 $B(n,p)$ に従うとき

$E(X)=np$　　　$V(X)=npq$　　　$\sigma(X)=\sqrt{npq}$

本日のテーマ「正規分布による確率の計算」

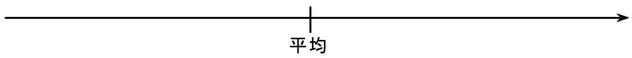
【問題①】
アラビア単語テスト１００問を解いてみよう。
あなたのクラスの結果を下のグラフにヒストグラムで表してみよう



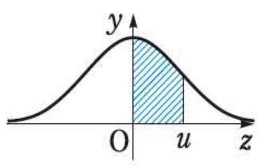
実は，この単語テストの正解数は，二項分布 $\boxed{}$ に従っています。
この単語テストの正解数の，平均・分散・標準偏差を求めよ。

正規分布曲線

サンプル数が多い二項分布は $\boxed{}$ に従うことが分かっています。
これは，確率全体を連続的に視覚化してくれる大変便利なものです。



正規分布曲線は，あまりにも有名なグラフでして，教科書数学Ｂの巻末の表を使えば，確率をすぐに計算することができます。
ただし，メモリの基準が標準偏差 σ です。

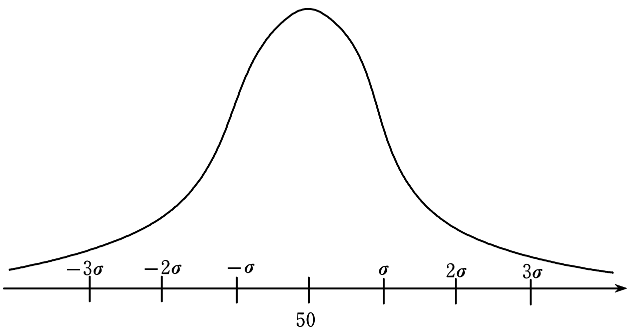


u	0	1	2
0.0	.0000	.0040	.0080
0.1	.0398	.0438	.0478
0.2	.0793	.0832	.0871
0.3	.1179	.1217	.1255
0.4	.1554	.1591	.1628
0.5	.1915	.1950	.1985
0.6	.2257	.2291	.2324
0.7	.2580	.2611	.2642
0.8	.2881	.2910	.2939
0.9	.3159	.3186	.3212
1.0	.3413	.3438	.3461
1.1	.3643	.3665	.3686

正規分布曲線を用いた確率の計算

＜例＞　アラビア単語100問テストで，60問以上得点する確率を求めよ。

プリント表の考察により平均 50，標準偏差 5 であるから



【問題②】アラビア単語100問テストで，35問以上65問以下の得点である確率を求めよ。

※ちなみに，この結果を３σの法則といいます。

＜例＞サイコロを 720 回投げたとき，１の目が 140 回以上出る確率を求めよ。

【問題③】3 つの中から 1 つ正解を選ぶ問題を全部で 72 問受けた。正解数が 20 問以下になる確率を求めよ。

※ここまでがノルマです

【おまけ問題】大学入試センター試験より

14000 人の生徒に対して，ある試験を実施した。試験の点数は正規分布に従うと考え，必要に応じて正規分布表を用いてもよい。
その試験の平均点は 66.2 点，標準偏差は 15.0 点であった。点数の高い順に順位をつけたとき，この試験で 80 点であった生徒の順位は $\boxed{\text{ア}}$ までにある。

$\boxed{\text{ア}}$ に当てはまるものを，次の ⓪～ⓓ のうちから一つ選べ。

⓪ 1 から 1000　① 1001 から 2000　② 2001 から 3000
③ 3001 から 4000　④ 4001 から 5000　⑤ 5001 から 6000
⑥ 6001 から 7000　⑦ 7001 から 8000　⑧ 8001 から 9000
⑨ 9001 から 10000　ⓐ 10001 から 11000　ⓑ 11001 から 12000
ⓒ 12001 から 13000　ⓓ 13001 から 14000

「統計学の基礎」第５回です。
いよいよ仮説検定です。確率を用いて，物事を判断する力を育成します。

＜これまでの復習＞

□下のような確率分布

X	x_1	x_2	………	x_n	計
確率	p_1	p_2	………	p_n	1

において，

■平均（期待値）　$m=E(X)$
$=x_1p_1+x_2p_2+x_3p_3+\cdots\cdots+x_np_n$
※「得点×確率」の和

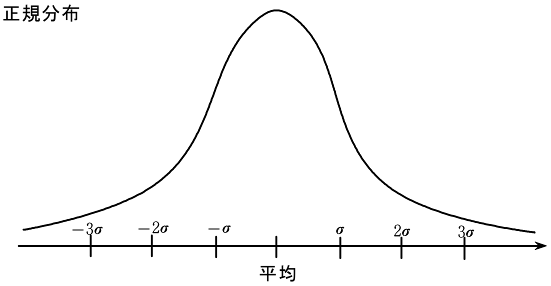
■分散　$V(X)=(x_1-m)^2p_1+(x_2-m)^2p_2+\cdots\cdots+(x_n-m)^2p_n$
※(得点－平均)^{2}の平均

■標準偏差　$\sigma(X)=\sqrt{V(X)}$　　※ 標準偏差＝$\sqrt{分散}$

□二項分布
「起こる」確率 p，「起こらない」確率 $q=1-p$　のとき，
n 回の試行のうち起こる回数を X としたときの確率分布のことを二項分布といい，$B(n,\ p)$ と表す。

確率変数 X が，二項分布 $B(n,p)$ に従うとき
$E(X)=np$　　$V(X)=npq$　　$\sigma(X)=\sqrt{npq}$

□正規分布



上のグラフのように，確率をグラフで表したものを正規分布曲線といいます。確率を面積で表します。教科書巻末の表を用いて確率を求めることができます。

【復習問題】
前回実施した「アラビア単語テスト１００問」について

この正解数を X とすると，X は二項分布 ＿＿＿＿ に従う。

X の平均は ＿＿＿＿

X の分散は ＿＿＿＿

X の標準偏差は ＿＿＿＿ である。

教科書巻末の正規分布表を用いて，この単語テストで５５点以上得点する確率を求めよ。

本日のテーマ「仮説検定の考え方」

【例題】
あるコインを 64 回投げたら，表が 41 回も出た。
このコインには偏りがあることを示したい。
有意水準５%で検定せよ。

このような疑問について検定する方法のことを仮説検定といいます。

＜仮説検定の手順＞
証明したいもの（コインに偏りがある）の正反対の仮説を考えます。これを帰無仮説といいます。これはいずれ棄却（否定）される予定です

帰無仮説 ＿＿＿＿

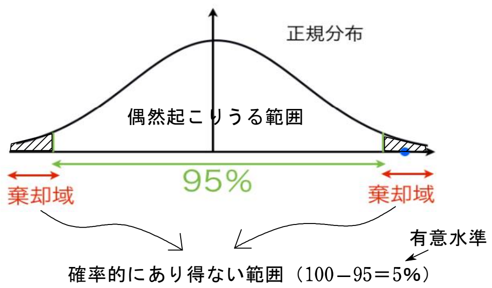
これを何とかして棄却（否定）すればよいのです（背理法っぽい）。

仮に，コインに偏りがなく，表裏の確率が $\frac{1}{2}$ であるとしよう
→ そんなコインを投げたら表が41回もでた。
→ そんなことは確率的にあり得ない！！
→ ということは，最初の仮説「偏りがない」が間違っているのだ
→ よって，帰無仮説は棄却され，このコインには偏りがあることが示された。

さて，上の文章の中に２か所アンダーラインがある場所があります。

１つは ＿＿＿＿，もう１つは ＿＿＿＿

この２つが仮説検定のミソです。



実際の確率が，このあり得ない範囲に入っていれば，仮説（偏りはない）は棄却され，「偏りがある」ことが示されることになります。

では，実際に解いてみましょう。

【例題】
あるコインを 64 回投げたら，表が 41 回も出た。
このコインには偏りがあることを示したい。
有意水準５％で検定せよ。

【解答】　帰無仮説を
「コインには偏りがなく，表の出る確率は $\frac{1}{2}$ である」とする。

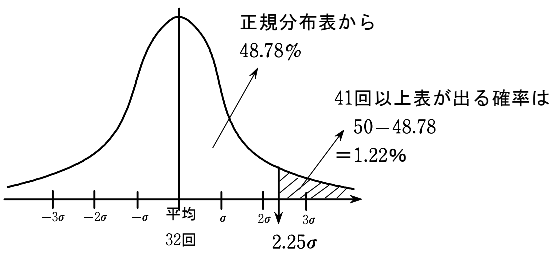
このとき，コインを 64 回投げたときの表の出る回数は，二項分布 $B\left(64,\ \frac{1}{2}\right)$ に従うので，

平均は $64\times\frac{1}{2}=32$，分散は $64\times\frac{1}{2}\times\frac{1}{2}=16$，標準偏差は $\sqrt{16}=4$

実験で，41 回表が出たということは，平均から 41－32＝9 回離れている。これを，標準偏差メモリに合わせるために，標準偏差で割ると

$9\div 4=2.25$

となり，標準偏差 2.25 個分平均から離れていることがわかる。



有意水準は５％だから，全体で５％未満の確率になれば「確率的にあり得ない」となる。
しかし，有意水準は右端左端合わせて５％だから，右半分だと 2.5％未満になれば「確率的にあり得ない」となる。
41回以上表が出る確率 1.22%であり，これは確率的にあり得ないと判断できる。
よって，「コインには偏りがなく，表が出る確率が $\frac{1}{2}$ である」という仮説は棄却され，コインには偏りがあると考えるのが妥当である。

注：もし有意水準を１％に設定すると，右半分で 0.5％未満にならないと「確率的にあり得ない」とは言えません。
今回は 41 回以上出る確率は 1.22％であるため，有意水準１％だと棄却できません。
このときは「偏りがない」ことが示されたわけではなく，「偏りがあるとまでは言えない（もう少し実験回数を増やせ）」という微妙な結論になります。

【問題】
あるサイコロを 180 回投げたら，１の目が 41 回出た。
このサイコロはいかさまサイコロであることを示したい。
有意水準５％で検定せよ。

【解答】

統計学の基礎の授業お疲れさまでした。
統計学とは「確率をもとに判断すること」です。
本時の仮説検定は，まさにその「判断」を主眼としたものです。
これをやってはじめて統計学をかじったと言えそうです。
最後に感想をお願いします。

おさらいとして，これまでの学習内容の要約を記します。
第１回　データの分析と標本調査の復習
第２回　確率における期待値と分散の学習
第３回　二項分布の学習　〇×二択クイズの確率。
第４回　アラビア単語テストの確率を，正規分布表から分析
第５回（本時）　仮説検定　コインやサイコロの偏りの検定

【感想や分かったこと】
評価基準　①左の問題が解けている　②感想を書いている

評価