2005. 7. 22

第3単元

離散量に関する推定と検定

目次

第3単元
離散量に関する推定と検定

単元のねらい　第1単元，第2単元では得られたデータをいかにまとめ，それからどのような情報を抽出したら良いかを学んだ．これは統計学の中でも 記述統計学 といわれる領域である．第3単元と第4単元では，統計学のもう一つの大きな領域である推測統計学 について学習する．推測統計学の中心は検定と推定であり，第3単元では離散量を扱う離散分布の検定と推定，第4単元では連続量を扱う連続分布の検定と推定の諸方法を学ぶ．

まず，§1 ではコインの例を用いて検定の基本的な考え方を身につける．§2 では，離散分布の代表的分布である2項分布を学び，割合に関する検定の考え方を導く．§3 では割合の区間推定，割合の差の検定と区間推定を学習する．§4 では，2項分布以外の多項分布，ポアッソン分布，幾何分布を取り上げ，最後にこの単元で学んだ諸手法の関係を明らかにする．

1　仮説検定の基礎

1.1　簡単な例による問題提起

(1)　3つの例

例1　大変むずかしい手術で成功率は10%といわれている．ある病院ではこの手術をこれまで9回実施したが，すべて失敗した．10人目の患者にこの手術を適用するときに，医師は患者に以上の説明をした後に，「あなたは幸運です，今度は成功する順番です」といった．

これは，寓話（笑い話）としては面白いが，誰も信用しないであろう．

例2　10人で あみだ をした．作ったくじは，1本は当たりで，残りの9本は外れである．9人がくじを引いたが全員が外れであった．最後に残った一人は，くじを引く前に，「私は当たりだ」と飛び上がって喜んだ．

これは，正しいようである．

それでは，上にあげた2つの例の基本的な違いはどこにあるのだろうか．

第1単元で，母集団とサンプルについて学んだ．これを上の2つの例に当てはめると，

　例2は，10本のあみだくじが母集団で，母集団全部について判断しているのに対して

　例1は，過去にこの手術を実施した患者の集団（数千人）が母集団で，母集団での成功率が10% であった．この病院に来た10人の患者はサンプルであるということになる．

この極めて単純な例から，「サンプルについての観測値から結論を出すのは，それほど単純ではない」ということが理解できるであろう．

例3　成功率が高い手術がある．ある病院ではこれまで100例の手術を実施し，失敗は1度もなく，成功率は100%であった．これから，患者に「この手術は完璧で，失敗する危険は全くありません」ということができないことは，これまで統計を学んだ皆さんにとっては，自明であろう．

それでは，「100回の内に1回もなかったのであるから，失敗率は1% 以下です」といえるであろうか．

この問題は，「製品ロットから100個のサンプルを取り出して検査したが，1個も不良品が見つからなかった」という場合に拡張することができる．

(2)　仮説検定

例3 手術の話で考えてみよう．手術の本当の失敗率 π が 0.01 で，$n = 100$ 人の患者に手術をしたとき，失敗数 x が 0 となる（100例全部が成功する）確率 p_{100} はいくらになるであろうか．

1人の患者に成功する確率は $p_1 = 1 - \pi = 1 - 0.01 = 0.99$ である．2人の患者に手術したとき，両方が成功する確率は $p_2 = (1 - \pi)^2 = 0.9801$ である[1]．

これを $n = 100$ 人に拡張すると，全員が成功する確率は次の式で求められる．

$$p_{100} = (1 - \pi)^n = 0.99^{100} = 0.366 \tag{1.1}$$

これより，失敗率 π が 0.01 であっても，「100人の患者全員に成功する」という結果はしばしば（約1/3の確率で）起こることが分かる．したがって，このことから，$\pi \leq 0.01$ であるとはいえない．

このように，100人で全員に成功したという事実から，失敗率が1%以下であるといえるかどうかを客観的に評価する方法を **仮説検定** といい，次の節以降で説明する．

[1] この式は，確率の基本定理である「独立事象の乗法定理」から導かれる．確率の基本定理については，§1.6 補足で説明する．

(3) 区間推定

前項で，100 人全員に成功しても，失敗率が $\pi = 0.01$ 以下であるということはできなかった．

それでは，失敗率はいくら以下であるというためには，どうしたら良いかを考える．

上に示したように，失敗率が 0.01 でも 100 人全員が成功ということがしばしば（約 1/3）起こる．それでは，失敗率 π が 0.03 のとき，100 人全員が成功する確率はいくらになるであろうか．それは，$p_{100} = (1-0.03)^{100} = 0.97^{100} = 0.048$ となる．失敗率 π が 0（必ず成功する）でないかぎり，p_{100} は 0 にならない．

そこで，上に計算した p_{100} がある基準値よりも小さいとき，偶然に起こったとは考えにくいと判断することにする．基準値としていくらを取るかは，問題によっていろいろと判断する必要があるが，一般には 0.05 が用いられる．

上に計算した確率 0.048 は 0.05 よりも小さい．これから，失敗率 π が 0.03 のとき「100 人で無失敗」ということは起こりにくいから，「失敗率は 0.03 未満であろう」という結論を出すことができる．

この考え方は，区間推定 といわれ，§3.1 で詳しく取り上げられる．

本日のまとめ

今日の学習で，非常に簡単な例によって，仮説検定と区間推定とはどんなものか，それらはどのような考えで導かれたかが漠然と理解できたであろう．

今日の学習の中で，偶然とは考えない基準として一般には 0.05 が用いられると説明した．この点については明日さらに詳しく取り上げる．

受講生は，ここで用いた例を，各人の身近な例に置き換えて，考えの筋道を理解してほしい．

1.2　賭けに不正はないか

17 世紀にフランスの確率論は,「とばく」の世界から発展した．そこで，ここでも，賭けの話から始めることにしよう．

サイコロを投げ偶数の目が出れば勝ち，奇数の目が出れば負けという賭けを考える．サイコロの代わりに，コインを投げて，表であれば勝ち，裏であれば負けとしても本質的には同じである[2]．

あなたは，このような賭けに参加し，4 回負け続けて，毎回賭け金を失ったとする．このとき，あなたなら，どのような態度を取るであろうか．取りうる道は，次の 2 通りある．

(i)　「4 回とも負ける」などというのは，相手が「いかさま」をしているに違いないのだから,「賭け を止めて席を立つ」．

(ii)　勝敗は時の運である．4 回続けて負けることも，ついてないときには起こりうるのであるから，第 5 回目・第 6 回目には勝つことを期待して,「さらに賭けを続ける」．

さて，どちらが正しい態度であろうか．真実は，相手が「いかさま」をしているか，いないか，のどちらかであるから，それが分かるものなら，何も迷うことはない．しかし，現実には勝った，負けたの事実の情報しかない．それでは，どのような統計的推論を導くべきであろうか．

(1)　一般人の直感的判断

ある先生が，大学の統計入門の講義に先立って，学生を対象として

「不正がなければ勝率が 50% である賭けで，
何回負け続けたら，いかさま と思うか」

という質問をし，表示 1.1 の結果が得られた．

6 回前後負け続けるとおかしいと思う人が多い．これはどんな基準で判断しているのであろうか？

---

[2] 通常考えられる賭けは，参加者が「奇数」,「表」と宣言し，当たれば勝ち，当たらなければ負けというものであろうが，ここでは，本文のような賭けを想定する．

表示1.1: 負け続ける回数に対する学生の判断の頻度

回数	2	3	4	5	6	7	8	9	計
人数	0	0	6	24	33	14	3	0	80

この問題について，考えてみよう.

賭けに不正がなければ（例えば，コインに細工がされていない）ことが確かであるものとすると，勝率は 0.5 (50%) となるはずである.

2 回賭けをして，2 回とも負ける確率は，$0.5 \times 0.5 = 0.5^2 = 0.25 = 25\%$ である.

賭けの回数を 3 回，4 回，... と増やしていくと，全部負ける確率は次のように急速に小さくなる.

表示1.2: 回数と，全敗および全敗＋全勝の確率 (%) との関係

回数	2	3	4	5	6	7	8	9
全敗の確率	25.0	12.5	6.25	3.12	1.56	0.78	0.39	0.20
全敗＋全勝	50.0	25.0	12.5	6.25	3.12	1.56	0.78	0.39

これと，学生の判断の結果を見比べると，数パーセント以下の確率でしか起こらないことを体験したとき，おかしいな と感じるようである.

これまでは，負け続けたとき,「おかしい，いかさま ではないか」と考えた.
それでは，勝ち続けたとき，どう考えるであろうか？

(i)　しめしめ，今日は運が良いと考え,「賭けを続ける」.

(ii)　甘い汁を吸わせて，賭けにのめり込ませる策略ではないかと疑い,「賭けを止める」.

の 2 つが考えられる.

前者は「楽天家」の考えで，相手の策略に乗ってしまう危険がある.

後者の場合，つまり，いかさまに引っかかる用心のためには,「全敗の確率」と共に「全勝の確率」も考えなければならない．その確率が，表示1.2 の「全

敗+全勝」の行に示されている．

このような常識的な判断に，根拠を与えてくれるのが，統計的検定 の考え方である．

(2)　問題

全敗，全勝という極端な場合ではなく,「負けが込んだ時どう判断するか」という問題を取り上げる．

コインを10回投げて，表が4回しか出なかった．これから，このコインは「表が出にくい」と結論して良いであろうか？

統計の基礎知識のない人でも，恐らく「これだけの実験からこの結論を出すには，根拠が弱いのではないか」と感じるであろう．

10回投げて表が全く出ないときどう考えるか．これなら，インチキ臭いと感じるであろう．それでは，表の回数が1回，2回，... のときどう考えたら良いであろうか？

(3)　実験

振れなどがない，新しいコインと，表示1.3のような表を準備する．

コインを投げて，表なら1，裏なら0を記入する．順次横に記録し，10回投げたならば，計の欄に1（表）の回数を記入する．

このような実験を100回繰り返す．

表示1.3: コイン実験（一部）

実験番号	1	2	3	4	5	6	7	8	9	10	計
1	0	1	0	1	1	0	0	0	1	0	4
2	0	0	0	0	0	0	1	0	1	0	2
3	0	0	1	0	1	1	1	0	0	0	4
4	0	1	1	0	1	1	0	1	0	0	5
5	1	0	1	0	1	0	1	1	0	1	6
6	0	1	1	1	1	1	0	1	0	0	6
7	1	1	1	0	1	0	0	1	0	1	6

実験は100回繰り返したが，表示1.3 には，最初の7回分だけを示す．

100回の実験結果を表示1.4 のようにまとめる．

表示1.4: 実験結果のまとめ

個数	度数	ヒストグラム	理論値(%)
0	0	|	0.10
1	1	|*	0.98
2	4	|****	4.39
3	14	|**************	11.72
4	13	|*************	20.51
5	26	|**************************	24.61
6	22	|**********************	20.51
7	11	|***********	11.72
8	8	|********	4.39
9	1	|*	0.98
10	0	|	0.10
合計	100		100.00

度数表を見ると，予想通り，5回が最も多く，大部分（90回/100回）は3回から7回である．

全部が表または裏ということは1度も起こらなかったが，実験回数を10倍の1000回に増やすと，起こるかもしれない．

全部が表または裏になる可能性が0でないとすると，表が全く出なくても,「運が悪かった」といわれれば，反論のしようがない．

(4) Excel による実験

表示1.3 の実験は，実際にコインを投げなくても，Excel を使って擬似的に体験できる．

表示1.5 にExcel シートの一部を示す．

Excel のトップメニューから「ツール」,「分析ツール」,「乱数発生」を選択する[3]．

[3] 「ツール」のプルダウンメニューに「分析ツール」がない場合は,「ツール」,「アドイン」を選択し，アドインリストから「分析ツール」の左をクリックする．

表示1.5: Excel **によるシミュレーション**

	A	B	C	D	E	F	G	H	I	J	K	L
1												0.5
2		1	2	3	4	5	6	7	8	9	10	個数
3	1	0.44	0.18	0.63	0.99	0.12	0.55	0.90	0.14	0.50	0.81	5
4	2	0.38	0.71	0.54	0.17	0.46	0.82	0.42	0.06	0.70	0.94	5
5	3	0.81	0.72	0.98	0.03	0.56	0.92	0.62	0.80	0.06	0.39	3
6	4	0.13	0.06	0.30	0.36	0.75	0.77	0.93	0.95	0.43	0.68	5
7	5	0.30	0.68	0.41	0.28	0.87	0.88	0.03	0.94	0.86	0.36	5
8	6	0.95	0.87	0.59	0.01	0.48	0.78	0.40	0.65	0.09	0.03	5
9	7	0.58	0.96	0.75	0.05	0.10	0.69	0.46	0.65	0.80	0.03	4
10	8	0.27	0.35	0.61	0.31	0.28	0.05	0.50	0.87	0.12	0.35	8
11	9	0.90	0.43	0.06	0.72	0.08	0.64	0.43	0.85	0.30	0.13	6
12	10	0.55	0.57	0.24	0.91	0.77	0.34	0.58	0.26	0.75	0.94	3

「乱数発生」のメニューの「分布」から,「均一」を選択する.「変数の数」を 10 **に,「乱数の数」を** 100 **に設定し,「出力先」に** B3 **を指定する.**

OK **をクリックすると**10**列,** 100**行の一様乱数(**0 **と** 1 **の間の乱数)が得られる.**

表頭に 1〜10**,表側に** 1〜100 **を入力する.**

一様乱数の値が 0.5 **以下であれば「表が出た」とすると,正しいコインを投げる実験と実質的に同じになる.**

横に並んだ10**個の乱数の中で** 0.5**以下の個数を数える.**

そのために, L1 **のセルに** 0.5 **を入力し,** L3 **のセルには,第**2**単元で学んだ** FREQUENCY **関数**

```
=FREQUENCY(B3:K3,$L$1)
```

を入力する.これを下にコピーする.

表示1.5**の** L **列から,第**2**単元の**§3.1 **で説明した方法で,表示**1.4 **の度数表とヒストグラムが作られる.**

表示1.5**の** L1 **の** 0.5 **を修正すると,表の出る確率が** 0.5 **ではなくなったときに,表の個数の分布がどのように変化するかを見ることができる.**

表示1.5の 0.5 の値（表の出る確率）を π で表わす[4].

0.5 を $0 < \pi < 1$ の範囲の任意の値に設定すると，表の出やすさを細工をしたコイン投げを体験できる.

表示1.4の 理論値の欄については，後に説明する.

■　演習1　　年末大売出しの抽選券に,「3枚につき1枚が当たる」と書いてある．この抽選券を6枚もらった．平均的には2枚当るので，少なくとも1枚は当たると思って抽選したところ，全部外れであった．この結果をあなたはどう判断するか？　根拠を述べなさい.

抽選券の枚数が 7, 8枚の場合はどうなるか？　■

本日のまとめ

前日に引き続き，仮説検定の本題に入るための準備であった.

Excel によるシミュレーションはぜひ自分で実行してほしい．この方法をマスターしておくと，将来いろいろの場面（テキストの内容がはっきりしない場合，自分のデータを解析する際，どの手法を取り上げたら良いか，など）で役に立つであろう.

1.3　統計的論理の基礎

(1)　命題と対偶(1)

今度は盗難事件を例として考える.

「H氏が犯人である」とすれば,「H氏が事件発生時に事件現場にいた」ことは確かである.

[4] 通常 π は円周率を表わすが，ここでは真の割合を表わすことにする．π は p に対応するギリシャ文字で，実際に得られた割合を p で表わすことに対応している.

この関係を

　「A（H 氏が犯人である）ならば

　　B（H 氏が事件時に事件現場にいる）が必ず成立する」

といい，$A \Rightarrow B$ で表わすことにする．

このとき，「B ならば A」（$A \Leftarrow B$ で表わす）が必ず成立するであろうか？ $A \Leftarrow B$ を具体的表現すると

　「H 氏が事件時に事件現場にいるとすれば，H 氏が犯人である」

となる．H 氏は犯人ではなく，たまたま事件現場に居合わせた可能性もあるから，必ずしも成立しないことが分かる．$A \Leftarrow B$ を $A \Rightarrow B$ の 逆 といい，「逆は必ずしも真ならず」ということになる．

それでは，「A でないとすれば B ではない」は正しいであろうか？　A でない，B でないを $\overline{A}$, $\overline{B}$ で表わすことにすると，$\overline{A} \Rightarrow \overline{B}$ は常に正しいであろうか．これは「H 氏が犯人でないならば，H 氏は事件時に事件現場にいない」ということになり，犯人でない人でも たまたま 事件現場に居合わせることがありうるので，常に成立するわけではない．これを，$A \Rightarrow B$ の 裏 といい，逆と同様「裏は必ずしも真ならず」である．

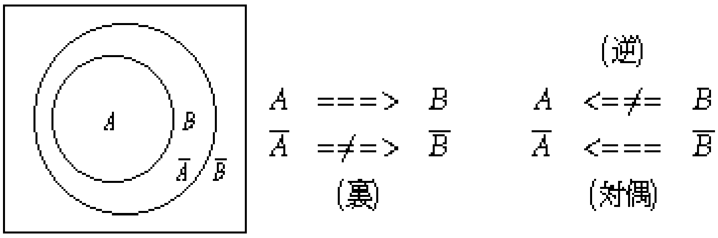
両方を組合わせた $\overline{A} \Leftarrow \overline{B}$ （Bでなければ，Aではない）はどうであろうか？「H 氏は事件時に事件現場にいないならば，H 氏が犯人ではない」は常に正しいことは理解できるであろう．これを 対偶 と呼ぶ．H 氏が事件現場以外にいたことが実証されると，アリバイありということで，無罪となるのは，この論理による．

逆 と 裏 は必ずしも成立しないが，対偶 は必ず成立することは，論理学の基本として，中学校でも教えられる．

上にあげた4つの関係は，表示1.6の右のように表わされ，A と B の関係は左の図で表わされる．

A の円は B の円の中に完全に含まれる．B の円の外側（$\overline{B}$）は A の円の外側（$\overline{A}$）に完全に含まれる．このような図は ベン図 と呼ばれる．

表示1.6: A と B の関係(1)



(2)　命題と対偶(2)

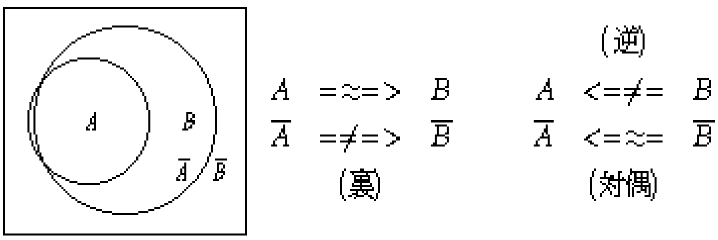
もし,「A ならば B である」が「A ならば ほぼ B である」となったらどうなるであろうか？

上の例でいえば, A :「H 氏が犯人である」===> B :「H 氏に動機がある」としたら, 必ずしも正しくない. しかし, おおむね正しい. これを A =≈=> B と表わすことにする.

この対偶「H 氏に動機がなければ, H 氏が犯人ではない」$\overline{A}$ <=≈= $\overline{B}$ もほぼ正しい.

この関係は, 表示1.7 のように表わされる.

表示1.7: A と B の関係(2)



表示1.6 では,「A の円は B の円の中に完全に含まれる」であったが, 表示1.7 では,「A の円は B の円の中にほぼ含まれる」となる.

表示1.7の関係を現実の場面で利用するためには,「ほぼ」というあいまいな表現では駄目である. 例えば,「A であれば 95% の確からしさで B である」というように, 定量的に表わす必要がある.

昔の日本では「動機なき殺人事件」は極めて稀であったから,「動機がなければ, かなりの確かさで犯人ではない」といえたが, 最近の日本では「動機が希薄な犯罪や殺人」が増えているので, この確かさはかなり低いであろう.

■　演習２　「A ならば B がほぼ成立する」という例を考え, その対偶もほぼ成立することを確かめよ.　■

(3)　賭けの問題への適用

A として「コインが正しい」, すなわち, 表が出る確率 π は 50% であるとする. もし, B として「10回中 表の回数は0回から10回の間である」とすれば, A $===>$ B であるが,「10回中 表の回数は2回から8回の間である」または「10回中 表の回数は3回から7回の間である」とすれば, A $=\approx=>$ B となる.

それでは, この2つの場合の ほぼ はどの程度の確かさであろうか？　表示1.4(p.9) の実験の結果を信用すると,

　B:「表の回数は2回から8回の間である」場合は 99/100
　B:「表の回数は3回から7回の間である」場合は 90/100

が ほぼ の程度を表わす数値となる.

しかし, この実験はわずか100回であるから, 心配である.

表示1.4 の理論値は, 実験ではなく, 理論的に求めた値である. これを使うと, 実験の結果 99/100, 90/100 に対応する値は 97.85%, 89.06% となる.

以下, 理論値の導き方を順を追って説明する.

(4)　表の回数の分布

前節のコイン投げ実験ではコインを10回投げたが, 説明を簡単にするために, まず, 3回投げる実験を考える.

コインは表か裏であるから, 3回の表裏の出方は全部で $2^3 = 8$ 通りの異なる

場合が起こりうる．それを列挙したのが表示1.8である．表であるときには ○ または 1，裏であるときには ● または 0 で示している．

表示1.8: 3回の実験で起こりうるすべての場合とその確率

		1回目	2回目	3回目	表の数(x)	組合わせの数	確率
1	●●●	0	0	0	0	1	1/8
2	○●●	1	0	0	1		
3	●○●	0	1	0	1		
4	●●○	0	0	1	1	3	3/8
5	○○●	1	1	0	2		
6	○●○	1	0	1	2		
7	●○○	0	1	1	2	3	3/8
8	○○○	1	1	1	3	1	1/8
計						8	8/8

コインに細工がされていなければ，ここに列挙した8つの場合はどれもが同じ確率で起こるはずである．

表が0回の場合は1通りであるが，表が1回は，第1, 2, 3回目に表が出る場合の3通りである．同様に，2回，3回の場合は，それぞれ，3, 1通りである．

この係数の並び 1, 3, 3, 1 は，2項係数と呼ばれ，異なる3個のものの中から 0，1，2，3個を選ぶときの 組合わせ (Combination) の数に等しい．

異なる n 個から x 個を選ぶ組合わせの数は ${}_nC_x$ で表わされ，

$$
\begin{aligned}
{}_nC_x &= \frac{n!}{x!(n-x)!} && (1.2)\\
&= \frac{n(n-1)(n-2)\cdots(x+1)x(x-1)(x-2)\cdots 2\cdot 1}{\{x(x-1)(x-2)\cdots 2\cdot 1\}\{(n-x)(n-x-1)(n-x-2)\cdots 2\cdot 1\}}\\
&= \frac{n(n-1)(n-2)\cdots(x+1)}{(n-x)(n-x-1)(n-x-2)\cdots 1} && (1.3)
\end{aligned}
$$

として求められる．ここに，$n!$ は n から 1 までの整数を掛け合わせたもので，n の 階乗と呼ばれる．ただし，$0!=1$ とする．

$n=3$ のとき，

$$
{}_3C_0 = \frac{3!}{0!3!} = \frac{3\cdot 2\cdot 1}{(1)(3\cdot 2\cdot 1)} = 1
$$

$$
\begin{aligned}
{}_3C_1 &= \frac{3!}{1!2!} = \frac{3\cdot 2\cdot 1}{(1)\,(2\cdot 1)} = 3 \\
{}_3C_2 &= \frac{3!}{2!1!} = \frac{3\cdot 2\cdot 1}{(2\cdot 1)\,(1)} = 3 \\
{}_3C_3 &= \frac{3!}{3!0!} = \frac{3\cdot 2\cdot 1}{(3\cdot 2\cdot 1)(1)} = 1
\end{aligned}
$$

である．これは，表示1.8の「組合わせの数」に対応している．

${}_nC_x$ を全組合わせの個数 $2^n = 2^3 = 8$ で割ると，表の回数が x 回になる確率が求められる．

ここで，$n = 10$ として，${}_{10}C_x$ を求め[5]，$2^n = 2^{10} = 1024$ で割ると，サイコロを10回投げたときの表の回数が x 回である確率が求められる．その確率を100倍して%で表わしたのが，表示1.4(p.9) の 理論値 の欄である．

本日のまとめ

命題と対偶の関係，論理的関係と統計的関係は，統計的仮説検定の基本であって，完全に理解されることが望まれる．

一般社会では,「A ならば B になるはずである」という命題から，B が表われると命題が証明されたという論理がかなり広く通用している．この論理は常に正しいものではないことを確認してほしい．

2項係数という，文科系の人には初めての式が出てきた．中学・高校で学んだ「順列・組合わせ」を思い出して，このような考え方に慣れてほしい．

1.4　仮説検定(1)

(1)　仮説検定とは

前に説明したコイン投げの賭けの例に戻る．

---

[5] 組合わせ数の性質と，Excel による求め方は §1.6(2) で説明する

以下の説明のために，正しいコインを $n = 10$ 回投げ，表の回数 x が $0 \sim 10$ の確率を，表示1.4 から再録する．

表示1.9: x 回の確率と累積確率 (%)

x	0	1	2	3	4	5	6	7	8	9	10
確率	0.10	0.98	4.39	11.72	20.51	24.61	20.51	11.72	4.39	0.98	0.10
累積	0.10	1.07	5.47	17.19	37.70	62.30	82.81	94.53	98.93	99.9	100
B			$\mid \Longleftarrow$			$98.93 - 1.07 = 97.86$			$\Longrightarrow \mid$		
B'				$\mid \Longleftarrow$		$94.53 - 5.47 = 89.06$		$\Longrightarrow \mid$			

「A: コインに細工がされていない」に対して

「B: コインを10回投げたとき，表の回数は 2回から8回の範囲内である」

としたとき，A =≈=> B であり，その「ほぼ正しい」の程度は，8回以下の確率から，2回未満（1回以下）の確率を引いて求められる．表示1.9 より $0.9893 - 0.0107 = 0.9785$，すなわち 97.85% である[6]．

また，

「$\overline{B}$: 表の回数が 0, 1, 9, 10」に対して

「$\overline{A}$: コインには細工がされている」

とすると，A =≈=> B の対偶 $\overline{A}$ <=≈= $\overline{B}$ も成り立ち，同様にその「ほぼ正しい」の程度は 97.85% となる．

「ほぼ 間違いがないであろう」というあいまいな表現ではなく，定量的に表わすと,「このような判断をしても間違える危険性は 2.15% に過ぎない」となる．

このような論理にもとづいて結論を導く方法を 仮説検定 という．

仮説検定では，はじめに A を定義する．この例では「コインに細工がされていない」または「表が出る確率は1/2 である ($\pi = 0.5$)」となる.「もし A が 正しければ」 という仮説である．これは，結論として，A ではない，すなわち，$\overline{A}$（コインには細工がされている）と否定される運命にある仮説である．そこ

[6] 表示1.9の 98.93, 1.07 には四捨五入の誤差が入っている．本文の 97.85% はもっと桁数の多い値から計算したので，最後の桁に1の違いが生じている．

で，A を **帰無仮説** (Null Hypothesis) （または，ゼロ仮説）と呼び，H_0 で表わすことにする．

結果が B に含まれないとき，すなわち，$\overline{B}$ であるとき，帰無仮説が否定される．否定されるを **棄却される** (reject される）といい，$\overline{B}$ を **棄却域** と呼び，一般には Reject の頭文字を取って R で表わされる..

帰無仮説が正しい場合でも結果が棄却域に入る確率は 0 ではない．すなわち，仮説検定では **誤りを犯す危険がある**．**危険率** は 100%と **ほぼ** との差であり，これを **第１種の誤りの確率** と呼び，その大きさを α で表わす．この例では $\alpha = 0.0215\,(= 2.15\%)$ となる．

結果が棄却域に入ったとき，帰無仮説が棄却され，$\overline{A}$ であると結論される．この $\overline{A}$ は **対立仮説** と呼ばれ，H_0 に対して H_1 で表わされる．

(2)　第１種の誤りの確率 α をいくらに取れば良いか

B を「表の回数は 2 回から 8 回の範囲内である」としたとき，ほぼの程度は $1-\alpha = 0.9785(= 97.85\%)$ で，第 1 種の誤りの α は 0.0215(=2.15%) であることが分かった．

もし，B を「表の回数は 3 回から 7 回の範囲内である」とすると，ほぼの程度は 89.06% で，第 1 種の誤りは 10.94% となる．

10 回に 1 回の割合で間違った結論を出すのは好ましくないと考えると，この B を採用することはできない．

この章の最初に述べたように，直感的な判断は，「数パーセントの誤りは許容しよう」ということであった．一般的には，第 1 種の誤りの確率として $\alpha = 0.05$ が用いられ，特に 第 1 種の誤りを犯したときの損失が大きいときには，$\alpha = 0.01$ を用いることがある．

コインには細工がされていないという帰無仮説 H_0 : 表の出る確率 $\pi = 0.5$ が正しいとき，起こりうる確率が 5%以下の現象が起こった場合に，表が出る確率 π が 0.5 とは違う（$\pi \neq 0.5$）と結論する．

これまでは 表の回数 x で考えてきたが，表の出た割合 $p = x/n$ で考えることもできる．

上の結論を，統計用語を用いて，

「表の出た割合 p と 帰無仮説の割合 $\pi = 0.5$ との間には有意な差がある」，

または

「表の出た割合 p と 帰無仮説の割合 0.5 との差は統計的に有意である」

と表現する．

もし，起こりうる確率が1%以下の結果が起こったときは,「差は統計的に高度に有意である」として区別することがある．

仮説検定において，この $\alpha = 0.05(5\%)$ または $\alpha = 0.01(1\%)$ を 有意水準 と呼び，それぞれの有意水準で有意になる値の肩に $*$ 印，または，$**$ 印をつけて示すことがある．

第1種の誤りの確率をもっと小さくしたいならば，すべて表 または すべて裏のときのみ帰無仮説を棄却すれば良い．そうすれば $\alpha = 2/1024 \approx 0.2\%$ となる．

この基準によれば，ずっと信頼のおける結論が出せることはいうまでもない．しかし，それは，いかさまの賭けを10回も続け，10回分の賭け金を失うという犠牲の上で達せられる．すなわち,「帰無仮説が誤りであるにもかかわらず，これを否定しない（これを見逃してしまう）誤り」がともなう．この誤りを，第2種の誤り といい，それをおかす確率を β で表わすことがある．

品質管理の分野では，第1種の誤りを あわて者の誤り， 第2種の誤りを ぼんやり者の誤り という．最初の文字「あ」がアルファに,「ぼ」がベータに対応しており，現場の作業者の理解を得るための工夫である．

第2種の誤りについては §3.3 で取り上げ，詳しく説明する．

また，α を 5% のように小さい値とすることが不適切な場合もあるので，どのような場で仮説検定を用いるかを配慮する必要がある．

本日のまとめ

今日は仮説検定とは何かという基本的な考え方が示され，その中で用いられる専門用語として，帰無仮説，対立仮説，棄却域，第1種・第2種の誤りとその確率 α, β がでてきた．初めて接する人にとって，完全に理解するにはかなり

の労力と努力が必要だろうと思う．理解が不十分だと思われる人は，明日の学習後にもう一度読み返すと，理解が深まるだろう．

1.5　仮説検定 (2)

(1)　両側検定と片側検定

賭け で，

i)　表の出た割合が 0.5 から非常に外れたときコインを疑う

ii)　表の出た割合が非常に小さくなったときコインを疑う

の2つの考えがある．

i) の場合には，前に述べたように，棄却域 $\overline{B}$ として，10 回のうちの勝数が $0,\ 1$ と $9,\ 10$ が取られ，$\alpha = 0.0215$ となる．

ii) の場合には，棄却域 $\overline{B}$ として，勝数が $0,\ 1,\ 2$ を取れば良い．そのとき，$\alpha = 0.0547$ で 5% をわずかに超える．

このように，棄却域を，両側に取る場合と，片側に取る場合があり，それぞれ，両側検定，片側検定 として区別される．

どちらを使うかは，事前に，有意であるという結論が出たとき，どのような行動を取るかを考えて決める．

「負け過ぎたときは賭けを止めるが，そうでなければ賭けを続ける」というのであれば，片側検定を採用する．それに対して，「負け過ぎたときだけでなく，勝ち過ぎたときも賭けを止める」ときは，両側検定が用いられる．

現実的な例を使って，両側検定と片側検定の使い分けを説明する．

工場で，新しい原料を使って，試作をした．

新しい原料から作った製品と従来の原料から作った製品の品質の差を検定すると，次の表の表頭に示す3つのいずれかの結果が得られる．

<table>
<tr><th>固有技術的判断</th><th colspan="3">統計的判断</th></tr>
<tr><th>対策</th><th>新しい原料が
有意に良い</th><th>両者に
有意差がない</th><th>従来の原料が
有意に良い</th></tr>
<tr><td>(1)</td><td colspan="2">新しい原料に切り替える</td><td>元の原料に戻す</td></tr>
<tr><td>(2)</td><td>新しい原料に切り替える</td><td colspan="2">元の原料に戻す</td></tr>
<tr><td>(3)</td><td>新しい原料に切り替える</td><td>試作を継続する</td><td>元の原料に戻す</td></tr>
</table>

どのような対策を取るかを3つの場合に分けて考える.

対策(1)　新しい原料は価格が安い（または，扱いやすい，などの良さがある）とか，系列のメーカーから購入する原料であるので，なるべく新しい原料に切り替えたい．したがって，

o　「従来の原料が有意に良い」ときには，原料を元に戻す.

o　そうでなければ，新しい原料に切り替える.

とする．これは，片側検定である.

対策(2)　新しい原料に切り替えるには，設備の変更などが必要であり，積極的に切り替える理由はないので，なるべく現在の原料を継続して使いたい．したがって，

o　「新しい原料が有意に良い」ときには，新原料に切り替える.

o　そうでなければ，従来の原料を使い続ける.

とする．この場合も片側検定であるが，対立仮説の方向が逆である.

対策(3)　新原料と従来の原料のどちらを選ぶかについて，優先度が，まったくない．したがって，

o　「新しい原料が有意に良い」ときには，新原料に切り替える.

o　「従来の原料が有意に良い」ときには，原料を元に戻す.

o　いずれでもない（「両者に有意差がない」）ときには，様子をみる.

という処置を取ることになる．これは，両側検定となる.

なお，両側検定で「様子をみる」の具体的な方法は，データを増やして再度検定したり，品質以外の要素を考慮してきめる，などがある.

この例のように，帰無仮説より，大きいとき，小さいとき，どちらでもないとき，の3通りに対して，どのような対応を取るかを明らかにすることにより，

片側検定，両側検定のいずれを用いるかが決まる．また，片側検定の場合は，棄却域をどちらに取るかが決まる．

■　演習３　上の対策(1)と対策(2)で，第１種の誤りと第２種の誤りは何か．α を小さくすると，どのような不都合が起こるか．■

(2)　仮説検定の限界

仮説検定によって，帰無仮説 A を証明することは論理的に不可能である．前にも述べたように $A \ <\!=\!\neq\!=\ B$ すなわち,「逆は必ずしも真ならず」であるから，表の回数が B の中に入ったからといって,「A コインが正しい」と結論することはできない．

本日のまとめ

昨日の仮説検定の基本をさらに発展させて，棄却域を片側に取る場合と両側に取る場合があることを，具体例を挙げて説明した．

また，最後に 仮説検定の限界 についても触れた．これは，仮説検定の誤用を避けるために大切な注意点である．

皆さんが直面する問題を取り上げて，片側検定・両側検定のいずれを用いるのが適当か，また，第１種・第２種の誤りがどのような意味を持つかを考えることによって，理解が深まるであろう．

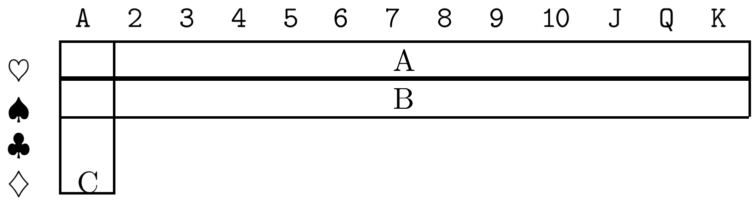
昨日と今日取り上げた事項は，来週以降の勉強の基本であるから，理解が不十分だと思われる人は，何回か読み返して，十分に理解を深めてほしい．

次の補足は，来週以降の理解には必須ではないが，確率・統計の教科書には必ず書かれているものである．勉強しておいて損はないであろう．

1.6　補足

(1)　確率の基本定理

ジョーカーを除いた 52 枚のトランプから 1 枚抜いたとき,「ハートである」という事柄を考える．確率論（確率に関する数学理論）では，確率を考えることのできる事柄を **事象** と呼ぶ．ここで,「ハートである」という事柄を事象 A ということにする．また,「スペードである」という事柄を事象 B とし,「エースである」という事柄を事象 C とする．



事象 A と事象 B の確率はいずれも 1/4，事象 C の確率は 1/13 であることは明らかである．これを $\Pr(A) = \Pr(B) = 1/4,\ \Pr(C) = 1/13$ と表わすことにする（Pr は Probability（確率）の意味である）．

事象 A と事象 B が同時に起こることはない．このように，同時に起こることのない事象を **背反事象** と呼び,「事象 A と事象 B は **互いに背反である**」という．一方，事象 A と事象 C あるいは 事象 B と事象 C は同時に起こる可能性があるので，背反事象ではない．背反事象の場合，そのいずれかが起こる事象の確率は，それぞれの事象の起こる確率の和として求められる．つまり,「ハートである，あるいは，スペードである（ハートまたはスペードである)」という事象（事象 $A+B$ と表わす）の確率は，それぞれの事象の確率の和として，

$$\Pr(A+B) = \Pr(A) + \Pr(B) = 1/4 + 1/4 = 2/4 = 1/2$$

のように求められる．このような関係を **背反事象の加法定理** という．背反事象の加法定理は，3 つ以上の背反事象の場合にも容易に拡張することができる．つまり，事象 $A_1, A_2, \ldots, A_n$ が互いに背反であるならば，

$$\Pr(A_1 + A_2 + \cdots + A_n) = \Pr(A_1) + \Pr(A_2) + \cdots + \Pr(A_n)$$

が成り立つ．

次に，事象 A と事象 C が同時に起こる，つまり「ハートであり，かつ，エースである」という事象（事象 AC と表わす）の確率を考える．まず，事象 A であることが分かったとする．そのとき，事象 C の確率は，ハートの中でエースである確率だから $1/13$ である．次に，事象 A でないことが分かったとする．そのとき，事象 C の確率は，「ハート以外」，つまり スペード，クラブ，ダイヤの中でエースである確率だから，やはり $3/39 = 1/13$ である．このように，事象 A であっても なくても 事象 C の確率が変わらないとき，両者を 独立事象と呼び，「事象 A と事象 C は 互いに独立である」という．事象 B と事象 C も互いに独立である．独立事象の場合，両事象が同時に起こる確率は，それぞれの事象の起こる確率の積として求められる．つまり，「ハートであり，かつ，エースである」という事象 AC の確率は，それぞれの事象の確率の積として，

$$\Pr(AC) = \Pr(A)\Pr(C) = 1/4 \cdot 1/13 = 1/52$$

のように求められる．このような関係を 独立事象の乗法定理 という．独立事象の乗法定理は，3 つ以上の独立事象の場合にも容易に拡張することができる．つまり，事象 A_1，A_2，$\cdots$，A_n が互いに独立であるならば，

$$\Pr(A_1 A_2 \cdots A_n) = \Pr(A_1) \cdot \Pr(A_2) \cdot \ \cdots \ \cdot \Pr(A_n)$$

が成り立つ．例えば，表の出る確率が $1/2$ のコインを続けて 5 回投げたとき，すべて表が出る確率は，1 回 1 回表が出る確率は $1/2$ で変わらないので，

$$1/2 \cdot 1/2 \cdot 1/2 \cdot 1/2 \cdot 1/2 = (1/2)^5 = 1/32$$

として求められる．

(2)　2 項係数

2 項分布の式に表われる 2 項係数 ${}_nC_x$ には面白い性質がある．

縦に n，横に x を取って 2 項係数を並べると表示1.10が得られる．

2 項係数は，左上と右上を足すことによって求められることが分かる．式で書くと，

表示1.10: 2 項係数の性質

n	
0	1
1	1 1
2	1 2 1
3	1 3 3 1
4	1 4 6 4 1
5	1 5 10 10 5 1
6	1 6 15 20 15 6 1
7	1 7 21 35 35 21 7 1
8	1 8 28 56 70 56 28 8 1
9	1 9 36 84 126 126 84 36 9 1
10	1 10 45 120 210 252 210 120 45 10 1

$$ {}_nC_x = {}_{n-1}C_{x-1} + {}_{n-1}C_x $$

となる．ただし，

$$ {}_nC_{-1} = 0 $$

とする．

表示1.10は パスカルの三角形 と呼ばれる．

$(x+y)^2,\ (x+y)^3$ を展開すると，$x^2+2xy+y^2,\ x^3+3x^2y+3xy^2+y^3$ となることは中学の代数で学んだであろう．

各項の係数 $1,\,2,\,1;\,1,\,3,\,3,\,1$ が 表示1.10 の $n=2,\ 3$ の行に対応している．

$x,\ y$ を $q=1-p,\ p$ とし，$2,\,3$ を n と一般化すると，$(q+p)^n$ となり，展開すると

$$ {}_nC_0\,q^n\,p^0 + {}_nC_1\,q^{n-1}\,p^1 + \cdots + {}_nC_x\,q^{n-x}\,p^x \\ + \cdots + {}_nC_{n-1}\,q^1\,p^{n-1} + {}_nC_n\,q^0\,p^n $$

となる．第 x 項は2項分布で x の確率になる．

$p+q=1$ であるから $(p+q)^n=1$，すなわち，確率の合計が 1 となることが証明できる．

2項係数 ${}_nC_x = \frac{n!}{x!(n-x)!}$ を，Excel で計算するには，階乗 (factorial !) を求める関数 FACT を用いて

```
=FACT(n)/(FACT(x)*FACT(n-x))
```

とするか，直接組合わせの個数 (combination) を求める関数 COMBIN を用いて

```
=COMBIN(n, x)
```

とする．

FACT 関数は引数が 170 を超えると桁あふれを起こして，結果が得られない．それに対して，COMBIN 関数は実用の範囲内 $n < 1000$ で解が得られる．

2　2項分布に関する推定と検定 (1)

2.1　2項分布 (1)

(1)　$\pi = 1/3,\ n = 3$ の場合

コインの問題は，表と裏が半々かどうかを問題としていた．言い換えれば，表である確率が $\pi = 0.5$ であるかどうかが問題であった．

確率が 0.5 ではない場合を考える．

表の出る確率 π が $1/3$ になるように細工されたコインがあるものと仮定する．

このようなコインを3回 (n) 投げて，表の出る回数 (x) の確率 (p_x) の分布を考える[1]．表示1.8 (p.15) は表示2.1のように変わる．

表示2.1: $\pi = 1/3,\ n = 3$ のとき，x の確率

組合わせ	1回目	2回目	3回目	表の個数 x	組合わせの数	確率 (p_x)
1	0	0	0	0	1	$1 \times (2/3)^3 = 8/27$
2	1	0	0	1		
3	0	1	0	1		
4	0	0	1	1	3	$3 \times (1/3)(2/3)^2 = 12/27$
5	1	1	0	2		
6	1	0	1	2		
7	0	1	1	2	3	$3 \times (1/3)^2(2/3) = 6/27$
8	1	1	1	3	1	$1 \times (1/3)^3 = 1/27$
計					8	$27/27$

組合わせ 8 は，3回とも表が出た場合（$x = 3$）であるので，表の出る確率 $\pi = 1/3$ の3乗 $p_3 = (1/3)^3 = 1/27$ となる．

組合わせ 1 は，3回とも裏が出た場合（$x = 0$）であるので，裏の出る確率

[1] §1.1 では，添え字は n で，p_n は $x = 0$ の確率を表わした．今後は，本文に説明したように，添え字は x を表わす．

$1-\pi=2/3$ の3乗 $p_0=(2/3)^3=8/27$ となる．

組合わせ 2 から組合わせ 4 までは,「表が1回，裏が2回出た」場合（$x=1$）である．表が何番目に出たかによって，3つに分かれている．この確率は $\pi(1-\pi)^2$ に組合わせの数を掛けたもの，すなわち，$3\times(1/3)(2/3)^2=12/27$ となる．

組合わせ 5 から組合わせ 7 も同様に考えて導かれる．

これを一般化すると

ある事象の起こる確率が π であるとき，n 回の試行をする．

n 回の内 その事象が x 回起こる確率 p_x

は

$$p_x = {}_nC_x\pi^x(1-\pi)^{n-x} \tag{2.1}$$

で求められる．

このように，ある事象が起こる確率 π と 試行回数 n が与えられたとき，x と p_x との対応関係を表わす分布を **2項分布** という[2]．

この例では,「ある事象が起こる」とは「表が出る」を意味している．

(2)　Excelによる確率の計算

式(2.1) に従って電卓で確率を計算するのは大変面倒であるが，Excel にはこの確率を計算する関数 `BINOMDIST` が準備されており，極めて簡単に求めることができる．

表示2.2は，x, n, π を入力すると，x となる確率，x 以下となる確率（累積確率），x 以上となる確率（上側累積確率）が求められる計算表である[3]．

D4，E4 には

```
=BINOMDIST( x, n, π, オプション)
D4: =BINOMDIST(A4,B4,C4,FALSE)
```

---

[2] ここに挙げた例のように，コインの面が 表 か 裏 のいずれかである場合，すなわち，結果が2通りに限られているときに，2項分布が用いられる．「好き」,「嫌い」,「どちらでもない」のように，結果が3通り以上あるとき，後に説明する多項分布になる．

[3] 通常 累積確率は 下側累積確率をあらわすので，下側を省略することが多い．
表示2.2 では，上側累積確率を単に上側確率と略記している．

表示2.2: Excel による 2 項分布の確率の計算表

	A	B	C	D	E	F
3	x	n	π	確率	累積確率	上側確率
4	0	3	0.3333	0.2963	0.2963	1.0000
5	1	3	0.3333	0.4444	0.7407	0.7037
6	1	10	0.5	0.0098	0.0107	0.9990
7	8	10	0.5	0.0439	0.9893	0.0547

```
E4: =BINOMDIST(A4,B4,C4,TRUE)
F4: =IF(A4=0,1,1-BINOMDIST(A4-1,B4,C4,TRUE))
```

が入力されている．最後のオプションを FALSE とすると x となる確率，TRUE とすると x 以下の累積確率が求められる．

x 以上となる確率は 1 から $x-1$ 以下の累積確率を引いて求められる．$x=0$ のときの上側累積確率はこの方法では求められず，かならず 1 になる．そのために，上に示したように IF 関数を使って計算する．

表示2.2 の上の 2 行は，表示2.1 の確率を計算したものである．0.2963, 0.4444 は $8/27=0.2963$, $12/27=4/9=0.4444$ である．また，累積確率の欄の 0.7407 は，上の 2 つの確率の合計 $(8+12)/27=20/27=0.7407$ である．

表示1.9（p.17）の下に，$\pi=0.5$, $n=10$ のとき，$2 \leq x \leq 8$ の確率が 0.9786 であると書かれている．この値を計算したのが，表示2.2 の下の 2 行である．2 以上，8 以下の確率は，8 以下の累積確率 0.9893 から 1 以下の累積確率 0.0107 を引いて求められる．

■ **演習4** 最後の例で，$3 \leq x \leq 7$ の確率 0.8906 を求めよ． ■

■ **演習5** ある工場の出荷検査では，各ロットから 20 個の製品を抜き取って検査し，一つも不良品がなかったときのみ合格としている．この検査方式で，不良率 π が 10%のロットが合格する確率はいくらか．また，不良率 π が 15% のロットと，5% のロットについてはどうか． ■

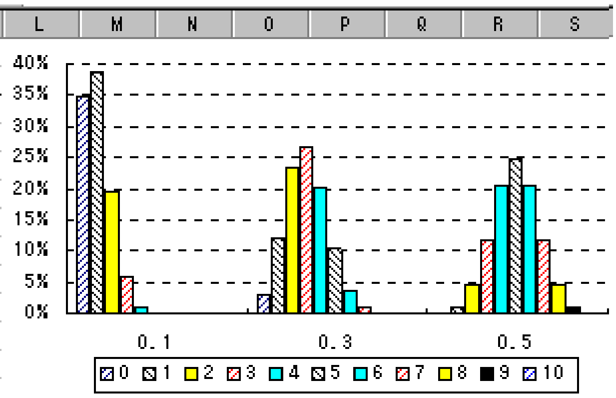
(3)　2項分布の形

前項で2項分布の確率を計算する方法を学んだ．それを使って，$n = 10$ の2項分布で，$\pi = 0.1$，0.3，0.5 と変化したときに，分布がどのように変化するかを調べる．

表示2.2 では，BINOMDIST 関数の3つのパラメータ x，n，π を横に並べたが，表示2.3 では，x を表側に n，π を表頭にとって計算表を作成する．

表示2.3: $n = 10$ の2項分布の確率とグラフ

	H	I	J	K
3	π	0.1	0.3	0.5
4	n	10	10	10
5	0	34.9%	2.8%	0.1%
6	1	**38.7%**	12.1%	1.0%
7	2	19.4%	23.3%	4.4%
8	3	5.7%	**26.7%**	11.7%
9	4	1.1%	20.0%	20.5%
10	5	0.1%	10.3%	**24.6%**
11	6	0.0%	3.7%	20.5%
12	7	0.0%	0.9%	11.7%
13	8	0.0%	0.1%	4.4%
14	9	0.0%	0.0%	1.0%
15	10	0.0%	0.0%	0.1%



I5 には，

```
=BINOMDIST($H5,I$4,I$3,FALSE)
```

が入力されている．このセルを下と右にコピーする．

確率の最大値は太字で表わしている．

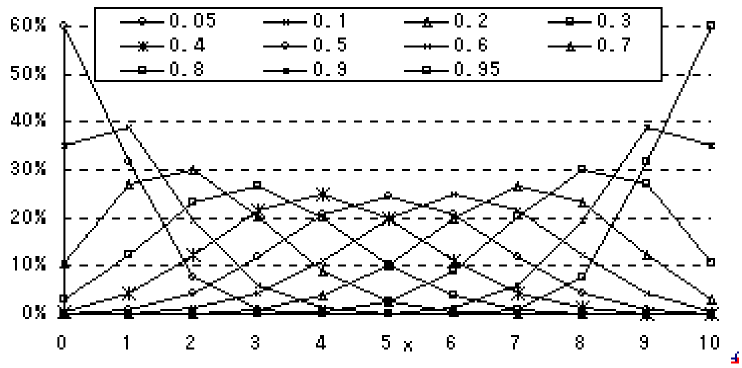
表示2.3 の右に3つの2項分布のグラフを示す．

$n = 10$ に固定し，π の値をもう少し小刻みに変化させて，π の値によって分布の形がどのように変化するかを見ることにする．

π による分布の形の変化は見にくいので，便宜的に表示2.4 のような折れ線で表わすことにする．

分布の最高値は，n と π の積になっていることが分かる．($\pi = 0.05$ のとき

表示2.4: π による 2 項分布の形の変化



$n\pi = 10 \times 0.05 = 0.5$ となり，最高値は $x = 0$ となる.）

$\pi = 0.5$ のとき，分布は左右対称であるが，π が 0.5 から外れるにつれて対称性が悪くなることが分かる. x は $0 \sim n$ の範囲に限られているため，$\pi = 0.0, 1.0$ に近いとき，外側に裾を引くことはできない.

■ **演習6** §1.2(4) の表示1.5（p.10）で，L1 のセルの値を変更すると，表示1.4 のヒストグラムが変化する.

得られたヒストグラムと表示2.4 と比較せよ. ■

次に，$\pi = 0.1$ を固定して，n を 10, 20, 40, 80 と変化させて，2 項分布の形がどのように変化するかを見たのが表示2.5 である.

表示2.4 のグラフは横軸に x が取られているが，表示2.5では，$p = x/n$ が取られている.

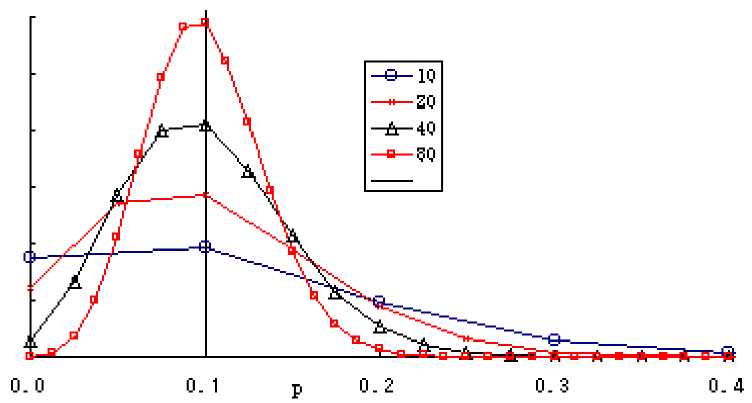
縦軸のスケールは，確率にn を乗じ，折れ線の下の面積がほぼ等しくなるように調整してある.

$\pi = 0.1$ の位置に縦の線が引かれている.

このグラフから次の性質を読み取ることができる.

- o $p = \pi$ の確率が最大になっている.
- o n が大きくなると，広がりが狭くなる.
- o n が大きくなると，左右対称に近づく.

表示2.5: $\pi = 0.1$ の 2 項分布のグラフ



本日のまとめ

今日は，コインが表か裏か，好きか嫌いか，良品か不良品か のように，yes, no のいずれかである場合，yes の個数 x がどのように変化するかを学び，Excel 関数で簡単に計算できることを知った．

表示2.2 の Excel による計算表を，下にコピーして，左の 3 つの値を入力することにより，2 項分布の確率や下側累積確率，上側累積確率が求められる．これから，２つの値の範囲内の確率などが自由に計算できるようにしておくと，今後でてくる Excel による計算表を自由に使いこなせるようになるであろう．

また，2 項分布の形が，yes の割合 π と 回数 n によってどのように変化するかをグラフで見た．

明日は，2 項分布 の特徴を数値で表わす「期待値」と「分散」を学ぶので，十分に理解しておいてほしい．

2.2　2項分布 (2)

(1)　x の期待値と分散

前節で，2項分布の位置や形は π, n によって変化することを見た．
このような分布の違いを数値で表わす方法を考える．

第1単元の §3.2 で学んだ期待値，分散と標準偏差を思いだし，$n = 3$, $\pi = 1/3$ の2項分布を例として，x の期待値と分散を求めてみよう．

表示2.6: 期待値と分散の計算

x	p_x	xp_x	$x - \mathrm{E}[x]$	$(x - \mathrm{E}[x])^2 p_x$
0	8/27	0/27	−1	8/27
1	12/27	12/27	0	0/27
2	6/27	12/27	1	6/27
3	1/27	3/27	2	4/27
合計	27/27	27/27		18/27
	=1	期待値		分散

表示2.6 の計算によって，x の期待値 $\mathrm{E}[x]$ と 分散 $\mathrm{V}[x]$ は

$$\mathrm{E}[x] = 27/27 = 1.0 \tag{2.2}$$

$$\mathrm{V}[x] = 18/27 = 2/3 \tag{2.3}$$

が得られる．

1/3の確率で起こる事象を3回繰り返せば，平均してその事象が1回起こると考えられるので，期待値は予想通りである．

分散については，直感的には理解できない．

そこで，最も単純な $n = 1$ の場合を考える．

$x = 0$ の確率は $p_0 = 1 - \pi = 2/3$，$x = 1$ の確率は $p_1 = \pi = 1/3$ であるから，x の期待値と分散は表示2.7 のように計算される．

すなわち，x の分散は $(x - \mathrm{E}[x])^2 p_x$ の欄の２つの値を合計し，

$$\begin{aligned}\mathrm{V}[x] &= (-\pi)^2(1-\pi) + (1-\pi)^2\pi \\ &= \pi(1-\pi)(\pi + (1-\pi)) = \pi(1-\pi)\end{aligned}$$

表示2.7: $n=1$ のときの期待値と分散の計算

x	p_x	xp_x	$x-\mathrm{E}[x]$	$(x-\mathrm{E}[x])^2 p_x$
0	$1-\pi$	0	$-\pi$	$(-\pi)^2(1-\pi)$
1	π	π	$1-\pi$	$(1-\pi)^2\pi$
合計	1	π		$\pi(1-\pi)$
		$\mathrm{E}[x]=1/3$		$\mathrm{V}[x]=2/9$

となる．

$n=3$ の x は，$n=1$ の試行を3回繰り返して得られた $x_1,\ x_2,\ x_3$ の合計であるから，その分散は，分散の加法性により，

$$\mathrm{V}[x]=\mathrm{V}[x_1]+\mathrm{V}[x_2]+\mathrm{V}[x_3]=3\times\pi(1-\pi)=2/3$$

となり，表示2.7の計算結果と一致する．

この式を一般化すると，2項分布に従う x の期待値と分散は

$$\mathrm{E}[x]=n\pi \tag{2.4}$$

$$\mathrm{V}[x]=n\pi(1-\pi) \tag{2.5}$$

で表わされる．

この式に $n=3,\ \pi=1/3$ を代入すると式(2.2), (2.3)が得られる．

(2)　p の期待値と分散

n 回の試行で事象の起こった割合を p で表わすことにする[4]．

$$p=\frac{x}{n} \tag{2.6}$$

π は理論的な割合（または全体での割合）であるのに対して，p は実際の賭けでの勝率，または n 人の調査で得られたテレビ番組の視聴率などである．

[4] x 回起こる確率を表わすにも同じ p の文字を使うので注意を要する．確率のときは，添え字として x，または，実際の数，0, 1, 2, ..., n が付く．

p は x を n で割った値であるから，p の期待値は x の期待値の $1/n$ であり，p の分散は x の分散の $1/n^2$ となる[5].

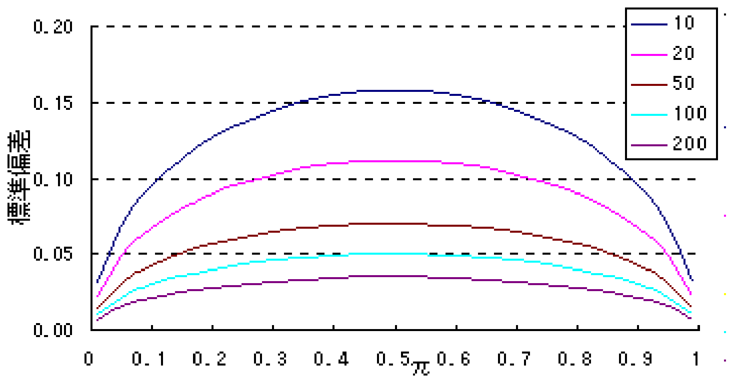
$$\mathrm{E}[p] = \pi \tag{2.7}$$
$$\mathrm{V}[p] = \frac{\pi(1-\pi)}{n} \tag{2.8}$$

なお，x, p の標準偏差は，上の分散の平方根である．

式(2.8) から，p の標準偏差は，n が大きくなれば小さくなり，また，同じ n でも π の値によって変化することが分かる．

この様子を表示2.8 に示す．

表示2.8: π, n による p の標準偏差の変化



表示2.8 は，横軸に π を取って，n ごとに p の標準偏差の変化の様子を示したものである．

$\pi = 0.5$ で最大になり，$\pi = 0, 1.0$ で 0 に近づくことが分かる．

また，n が小さいとき，標準偏差が相当大きいことを示している．例えば，$n = 10$, $p = 0.5$ のとき，標準偏差は約 0.17 となる．

$\pi = 0.5$ のときの標準偏差が 0.05 であるのはそれほど大きなバラツキではないが，$\pi = 0.1$ のときの標準偏差が0.05 であるのはかなり大きなバラツキで

[5] 第1単元の §3.3 分散の加法性の「性質2」で，$a = 1/n$ とすると求められる．

ある．

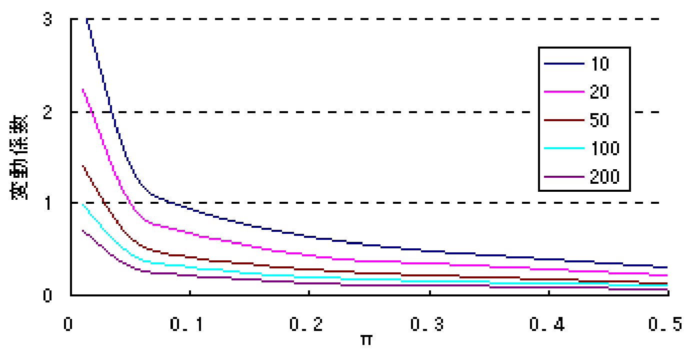
すなわち，バラツキの大きさの別の指標は，標準偏差が期待値に対してどのくらいの大きさをもっているかである．これは，

$$\text{変動係数} = \text{標準偏差}/\text{期待値}$$

で評価できる．

表示2.8の縦軸を標準偏差から変動係数に変えると表示2.9が得られる．

表示2.9: π, n による p の変動係数の変化



$\pi > 0.5$ に対しては，期待値が大きくなり，変動係数は誤解を与えるので，横軸の範囲から外した．その部分については，$1 - \pi$ に置き換えて考える．

これから，π が 0 に近づくと，p の標準偏差は小さくなるが，変動係数は急速に大きくなることが分かる．

これまでの数値例は π がかなり大きい場合であったが，もし，非常に起こりにくい（π が 0 に近い）事象について調べたいときには，直感的な予想以上にたくさんの観測が必要である（§3.3 (3) で再び取り上げる）．

本日のまとめ

多くの統計の教科書では，2項分布の期待値と分散の式(2.2), (2.3), (2.4), (2.5)が根拠抜きで示されている．このテキストでは，$n = 1$ の場合から，分散の加

法性で導いた．このように，公式を覚えるのではなく，自分で導くことができるようになると，応用が利き，公式を正しく利用することができる．

また，サンプルから求めた p の標準偏差は，n が小さいとき予想以上に大きいことを記憶に止めておいてほしい．

2.3　2項分布の正規分布近似

(1)　正規分布近似の方法

第1単元で，もとの分布によらず，複数個の和または平均の分布は正規分布に近づくことを学んだ．

2項分布も $n = 1$ の試行の結果を加えたものであると考えると，n がある程度の大きさがあれば，正規分布に近いであろう．

そこで，まず，$n = 10,\ \pi = 0.5$ の場合について調べる．

この2項分布の x の期待値と分散（標準偏差）は

$$
\begin{aligned}
\mathrm{E}[x] &= n\pi = 10 \times 0.5 = 5.0 \\
\mathrm{V}[x] &= n\pi(1-\pi) = 10 \times 0.5(1-0.5) = 2.5 = 1.581^2
\end{aligned}
$$

である．

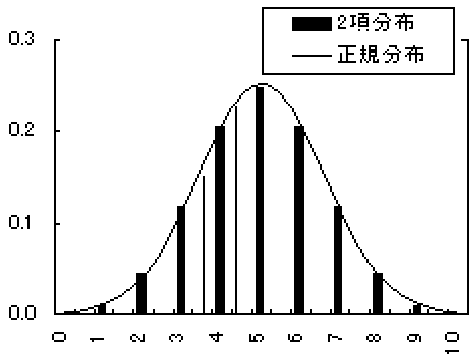
2項分布と正規分布近似との比較をするグラフは工夫が必要である．表示2.3(p.30) に示したように，2項分布は棒グラフで表わされる（表示2.4の折れ線は，便宜的なものであった）．それに対して正規分布は連続的に変化する分布で，曲線で表わされる．

そこで，2項分布の棒グラフに，平均が 5.0 で，分散が 2.5 の正規分布 N(5, 2.5) の曲線を重ねると表示2.10左 が得られる．

2つの分布は非常に良く合っているように見える．

2項分布で $x = 4$ の確率は 0.205 である（表示2.10）．

表示2.10: 2項分布（$n=10,\ \pi=0.5$）と正規分布近似



G	H	I	J
n	10		
π	0.50		
	2項分布	正規分布	差
0	0.0010	0.0022	0.0012
1	0.0098	0.0112	0.0014
2	0.0439	0.0435	-0.0005
3	0.1172	0.1145	-0.0027
4	0.2051	0.2045	-0.0006
5	0.2461	0.2482	0.0021
6	0.2051	0.2045	-0.0006
7	0.1172	0.1145	-0.0027
8	0.0439	0.0435	-0.0005
9	0.0098	0.0112	0.0014
10	0.0010	0.0020	0.0010
期待値	5.0000		
標準偏差	1.5811		

これに対応する正規分布近似の確率はどうしたら求められるのであろうか．表示2.10から，$x=4$ の棒の高さに対応する正規分布の確率は，$x=4\pm0.5$ の範囲の面積に対応する（この部分を，表示2.10では，縦線で示している）．

まず，$x=3.5$ と 4.5 が期待値から標準偏差の何倍離れているかを表わす偏差値 z を求める．

$$x=3.5,\quad z=\frac{3.5-5.0}{1.581}=-0.949$$
$$x=4.5,\quad z=\frac{4.5-5.0}{1.581}=-0.316$$

標準正規分布で上に求めた z の値よりも小さい確率は，Excel関数

```
=NORMSDIST(z)
```

で求められ，

$$\Pr(z<-0.949)=0.1714$$

$$\Pr(z<-0.316)=0.3759$$

となる．

これから，正規分布近似による $x=4$ の確率は

$$\Pr(x=4) \approx 0.3759 - 0.1714 = 0.2045$$

となる．この値は，2 項分布から求めた確率 0.2051 との差は 0.0006 である．

(2)　確率の近似の精度

以上の計算を，すべての x について計算した結果を表示2.10右に示す．

H6 には 2 項分布の確率が

```
=BINOMDIST($G6,$H$3,H$4,FALSE)
```

で計算されている．

正規分布近似の確率を計算するためには，$x \pm 0.5$ の偏差値が必要である．そのために，x の期待値 $n\pi$ と標準偏差 $\sqrt{n\pi(1-\pi)}$ が H17 と H18 に求められている．

I6 には $x \leq 0.5$ の確率が

```
=NORMSDIST(($G6+0.5-H$17)/H$18)
```

で計算されている．

I7 には $0.5 \leq x \leq 1.5$ の確率を求めるために，$\Pr(x \leq 1.5) - \Pr(x \leq 0.5)$ が

```
=NORMSDIST(($G7+0.5-H$17)/H$18)
  -NORMSDIST(($G7-0.5-H$17)/H$18)
```

で計算されている．

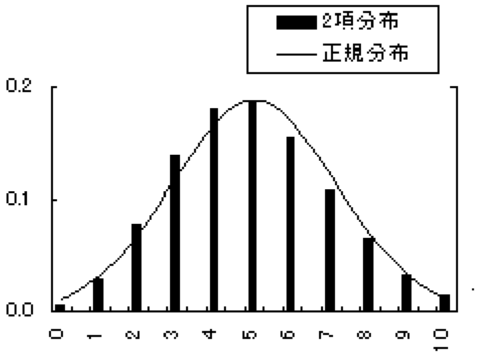
正確な確率と近似による確率の差は最大でも 0.0027 でかなり良く近似されていることが分かる．

$n=50$, $\pi=0.1$ の場合（x の期待値は前と同様に 5）に同様のグラフと計算した結果が表示2.11である．

この場合は，2 項分布が左右対称ではなく，$\pi=0.5$ の場合と異なり，グラフで見ても近似はあまり良くないことが分かり，数値的にも，差の最大値は 0.018 である．

表示2.4 (p.31) からも分かるように π が 0 または 1 に近づくと分布が左右対称ではなくなる．それを，左右対称の正規分布で近似するのであるから，当然の結果である．

表示2.11: 2項分布（$n = 50,\ \pi = 0.1$）と正規分布近似



n	50		
π	0.10		
	2項分布	正規分布	差
0	0.0052	0.0169	0.0118
1	0.0286	0.0325	0.0039
2	0.0779	0.0698	-0.0081
3	0.1386	0.1205	-0.0181
4	0.1809	0.1671	-0.0138
5	0.1849	0.1863	0.0014
6	0.1541	0.1671	0.0130
7	0.1076	0.1205	0.0128
8	0.0643	0.0698	0.0055
9	0.0333	0.0325	-0.0008
10	0.0152	0.0122	-0.0030
期待値	5.0000		
標準偏差	2.1213		

(3)　累積確率の正規近似

前項では，x がある値となる確率を近似的に求める場合の精度を確認した．

正規分布近似を使う場面では，x がある値である確率ではなく，x がある値以下（または 以上）である確率が多く用いられる．

表示2.10，表示2.11 と同じ条件で，2 項分布累積確率を求めたのが表示2.12 の「連続修正あり」の列である．

表の上半分は下側累積確率，下半分は上側累積確率である．

正規分布近似の下側累積確率と上側累積確率は次の関数で計算される．

```
N6:  =NORMSDIST((L6+0.5-M$17)/M$18)
N12: =1-NORMSDIST((L12-0.5-M$17)/M$18)
```

$x \pm 0.5$ に対する偏差値 z を求め，それから累積確率を計算している．

この 0.5 は，不連続分布の2項分布を連続分布の正規分布で近似するためのものであって，連続修正 と呼ばれる．

「連続修正なし」の列には，連続修正を加えない正規近似の結果が求められている．

表示2.12: 2 項分布の累積確率

	L	M	N	O	P	Q
3	n	10				
4	π	0.50	連続修正あり		連続修正なし	
5		2項分布	正規分布	差	正規分布	差
6	0	0.001	0.002	0.001	0.001	0.000
7	1	**0.011**	**0.013**	**0.003**	**0.006**	**-0.005**
8	2	**0.055**	**0.057**	**0.002**	**0.029**	**-0.026**
9	3	0.172	0.171	0.000	0.103	-0.069
10	4	0.377	0.376	-0.001	0.264	-0.113
11	5					
12	6	0.377	0.376	-0.001	0.264	-0.113
13	7	0.172	0.171	0.000	0.103	-0.069
14	8	**0.055**	**0.057**	**0.002**	**0.029**	**-0.026**
15	9	**0.011**	**0.013**	**0.003**	**0.006**	**-0.005**
16	10	0.001	0.002	0.001	0.001	0.000
17	期待値	5.000				
18	標準偏差	1.581				
19						
20	n	50				
21	π	0.1	連続修正あり		連続修正なし	
22		2項分布	正規分布	差	正規分布	差
23	0	**0.005**	**0.017**	**0.012**	**0.009**	**0.004**
24	1	**0.034**	**0.049**	**0.016**	**0.030**	**-0.004**
25	2	0.112	0.119	0.008	0.079	-0.033
26	3	0.250	0.240	-0.011	0.173	-0.077
27	4	0.431	0.407	-0.024	0.319	-0.113
28	5					
29	6	0.384	0.407	0.023	0.319	-0.065
30	7	0.230	0.240	0.010	0.173	-0.057
31	8	0.122	0.119	-0.003	0.079	-0.043
32	9	**0.058**	**0.049**	**-0.008**	**0.030**	**-0.028**
33	10	**0.025**	**0.017**	**-0.008**	**0.009**	**-0.015**
34	期待値	5.000				
35	標準偏差	2.121				

累積確率が用いられるのは，検定，区間推定のいずれの場合でも，その確率が 5% を中心とし，1% から 10% の範囲内であるから，その付近での近似の良さが大切である．

表示2.12 では，その範囲を太字で表わしている．

$\pi = 0.5$ で分布が対称のとき，累積確率の差の絶対値の最大は 0.003 で近似の良いことが分かる．連続修正を加えないと差の絶対値の最大は 0.026 と大き

くなる．

$n=50,\ \pi=0.1$ で分布が非対称の場合は，差の絶対値の最大は，0.016, 0.028 である．対称の場合に比べて大きくなる．

$n\pi=5$ を固定して，$n=100,\ 200,\ 1000$ と大きくした場合，および，$n\pi$ を 10, 20 にして同様の検討をした結果を次に示す．

表示 2.13: 正規分布近似による累積確率の精度

	連続修正あり			連続修正なし		
n	$n\pi=5$	$n\pi=10$	$n\pi=20$	$n\pi=5$	$n\pi=10$	$n\pi=20$
10	0.003			0.026		
20	0.011	0.001		0.021	0.021	
50	0.016	0.007	0.002	0.028	0.022	0.023
100	0.017	0.010	0.005	0.030	0.025	0.021
1000	0.018	0.012	0.008	0.031	0.026	0.022

これから，次のことが分かるであろう．

n が大きくなると，精度の値が小さくなる（精度が向上する）と考えられるが，上の結果を見ると，逆の傾向が見られる．

$n\pi$（x の期待値）が大きくなると，近似が良くなる．連続修正を加えると近似はかなり向上する．

したがって，x の値が小さいとき（目安は 5 以下）は，正規近似は使わないのが良いであろう．なお，π が 1 に近い（x が n に近い）とき，$n-x$ が小さいときも同様である．

$n\pi$ が小さい（正確にいえば，$n(1-\pi)$ が小さい場合も含む）場合には，正規分布の近似は悪くなる．しかし，連続修正を加えるとかなり良く合う．

本日のまとめ

今日は，2 項分布が正規分布で近似できることを学んだ．

正規分布については，第 4 単元で詳しく説明される．今日の内容について理解が不十分であったときは，第 4 単元を勉強してから，もう一度読み返すと，理

解ができるようになるであろう．

また，近似の精度が条件によって変化することを知った．

さらに理解を深めるために，表示2.12 の Excel ファイルで，$n\pi$，n を修正して，近似の良さの変化を見ると良いであろう．

ここで学んだ正規近似の性質は，次節以降で説明される仮説検定や区間推定で重要な役割を果たす．

2.4　仮説検定

(1)　正確法（棄却域を用いる仮説検定）

§1.4 では，仮説検定の基本的な考え方と手順を説明した．それは，

- 帰無仮説を設定する．
- 対立仮説を設定し，棄却域を求める．
- 観測された値が棄却域に含まれたら帰無仮説を棄却する．

であった．

帰無仮説 H_0：$\pi = 0.20$ を両側検定するために，$n = 100$ のサンプルを取った場合を考える．

棄却域を設定するためには，次の手順が必要である．

表示2.2(p.29) の π，n に0.20, 100 を入力する．

x に $n\pi = 100 \times 0.20 = 20$ を入力する．x を小さくすると，累積確率が減少する．累積確率が 0.025 以下になる x が下側の棄却域の上限値である．

$x = 12$ の累積確率は 0.0253 と 0.025 よりも大きい．$x = 11$ にすると，累積確率が 0.0126 と 0.025 よりも小さくなる．従って，下側の棄却域の上限値は 11 となる．

逆に x を 20 から大きい方から変化させる．$x = 28$ の上側累積確率は 0.0342 で，0.025 よりも大きい．$x = 29$ の上側累積確率は 0.0200 で，0.025 よりも

表示2.14: 2項分布による棄却域の計算表

x	n	π	確率	累積確率	上側確率
20	100	0.200	0.0993	0.5595	0.5398
12	100	0.200	0.0128	**0.0253**	0.9874
11	100	0.200	0.0069	**0.0126**	0.9943
28	100	0.200	0.0141	0.9800	**0.0342**
29	100	0.200	0.0088	0.9888	**0.0200**

小さくなる．従って，上側の棄却域の下限値は 29 となる．

これから，x が 11 以下，または，29 以上のとき，帰無仮説 $\pi = 0.2$ は棄却される．

片側検定の場合は，累積確率 または 上側累積確率 が 0.05 以下になる限界値を求めれば良い．

x を $n\pi$ から一つずつ変化させるのは煩雑である．試行錯誤で変化させることにより，解を求めるまでの手数を少なくすることができる．

■　演習 7　帰無仮説 $\pi = 0.30$ を $n = 50$ のデータで検定するための棄却域を，対立仮説が 両側・上側・下側 のそれぞれについて求めよ．ただし，$\alpha = 0.05$ とする．■

(2)　正確法（p値を用いる仮説検定）

前項の手順は伝統的なもので，昔からずっと使われてきた．

最近のコンピュータとプログラムの進歩によって，前節に説明した方法で，観測された値が帰無仮説が正しいときどのくらい起こりにくいかが簡単に求められるようになり，最近の統計解析プログラムでは別の手順を用いるのが主流になってきた．

上に用いた $n = 100,\ \pi = 0.20$ の場合を取り上げて説明しよう．

$x = 11$ という値が得られたとき，この値は前項の方法で求めた棄却域の範囲内に入っている．

新しい方法では，$n = 100$，$\pi = 0.20$ の2項分布で x が 11 以下の確率を計算する．これは，

```
=BINOMDIST(11,100,0.2,TRUE)
```

で計算され，0.0126 である．

このような考えで求められた確率 0.0126 を **p値** と呼ぶ．すなわち，帰無仮説が正しいとき，帰無仮説よりも 得られた値以上離れている値が得られる確率である．得られた値が得られる確率ではないことに注意．

p値が，予め決めておいた α よりも小さいとき，観測された値は棄却域に含まれると判断し，$\pi = 0.20$ と有意差があると結論する．つまり，帰無仮説 $\pi = 0.20$ を棄却する．

もし，両側検定を使う場合には，反対側の確率も含めなければならないので2倍し，0.0251 となる．

Excelを使ってこの手順の仮説検定を実行するためには，表示2.15 のような計算表を準備しておくのが良いであろう．

表示2.15: p値の計算表

	A	B	C	D	E	F	G	H	I
15					2項分布			正規近似	
16	π	n	x	下側	上側	両側	下側	上側	両側
17	0.2	100	11	0.0126	0.9943	0.0251	0.0168	0.9912	0.0336
18	0.2	100	12	0.0253	0.9874	0.0507	0.0304	0.9832	0.0608
19	0.2	100	29	0.9888	0.0200	0.0400	0.9912	0.0168	0.0336
20	0.5	100	39	0.0176	0.9895	0.0352	0.0179	0.9893	0.0357
21									
22	D17:	=BINOMDIST(C17,B17,A17,TRUE)							
23	E17:	=IF(C17=0,1,1-BINOMDIST(C17-1,B17,A17,TRUE))							
24	F17:	=2*MIN(D17:E17)							
25	G17:	=NORMSDIST((C17+0.5-B17*A17)/SQRT(B17*A17*(1-A17)))							
26	H17:	=1-NORMSDIST((C17-0.5-B17*A17)/SQRT(B17*A17*(1-A17)))							
27	I17:	=2*MIN(G17:H17)							

この計算表は，片側検定（下側，上側のどちら側でも），両側検定 のいずれにも対応できるように作られている．

左の3つのセルにデータを入力する．その右の3列に片側検定と両側検定のp値が求められる（右の3列については，次の項で説明する）．

片側検定のp値は，p 以下または以上の確率で，両側検定のp値は，下側確率と上側確率の小さい方の2倍である．

■　演習8　　演習7の検定で，$x = 19$ が得られたとき，両側p値と，上側p値を求めよ．　■

(3)　正規近似（棄却域を用いる仮説検定）

正確法で棄却域を決めるには試行錯誤が必要で実務的にはかなり煩雑である．

そこで，§2.3 で学んだ 正規近似を使って棄却域を設定する．

帰無仮説が正しいとき，x の期待値 $\mathrm{E}[x]$ と標準偏差 $\mathrm{D}[x]$ は，

$$\mathrm{E}[x] = n\pi = 100 \times 0.2 = 20, \quad \mathrm{D}[x] = \sqrt{100 \times 0.2(1-0.2)} = 4.0$$

である．

標準正規分布で，± 1.96 の範囲外の確率は 5% である（第1単元 §3.3 参照）．したがって，帰無仮説が正しければ，x は

$$\begin{gathered} \mathrm{E}[x] - 1.96\mathrm{D}[x] - 0.5 < x < \mathrm{E}[x] + 1.96\mathrm{D}[x] + 0.5 \\ n\pi - 1.96\sqrt{n\pi(1-\pi)} - 0.5 < x < n\pi + 1.96\sqrt{n\pi(1-\pi)} + 0.5 \end{gathered} \tag{2.9}$$

の範囲内に含まれる確率が95%である．すなわち，式(2.9) の範囲外が仮説検定の棄却域 B に対応する．

式(2.9) を

$$x \sim n\pi \pm (1.96\sqrt{n\pi(1-\pi)} + 0.5)$$

のように省略して表わすことにする．

上の例について計算すると，

$$\begin{aligned} x \sim&\quad 20 \pm (1.96 \times 4.0 + 0.5) \\ &\sim 20 \pm (7.84 + 0.5) \\ &\sim (11.66,\ 28.34) \end{aligned}$$

の外側が棄却域となる．x は整数値しか取らないから，下限値は切捨て，上限値は切上げて，$x \le 11,\ 29 \le x$ が棄却域となる．

この値は，前項の正確な棄却限界値 11, 29 と同じである．

片側検定の場合は，上の式の 1.96 の代わりに，1.645（片側5%点）を用いれば良い．

この方法により棄却域を求める計算表を表示2.16に示す．

表示 2.16: 正規近似による棄却域の計算表

	A	B	C	D	E	F	G	H	I
31	π	n	片側α	期待値	標準偏差	下限	上限	下限	上限
32	0.2	100	0.025	20.0	4.00	11.66	28.34	11	29
33	0.2	100	0.05	20.0	4.00	12.92	27.08	12	28
34									
35	D32:	=A32*B32				H32:	=ROUNDDOWN(F32,0)		
36	E32:	=SQRT(B32*A32*(1-A32))				I32:	=ROUNDUP(G32,0)		
37	F32:	=D32-NORMSINV(1-C32)*E32-0.5							
38	G32:	=D32+NORMSINV(1-C32)*E32+0.5							

両側検定の場合は，13 行目に示すように，$\alpha/2 = 0.025$ を入力する．

片側検定の場合は，14 行目に示すように，$\alpha = 0.05$ を入力し，対立仮説の方向により，12 以下 または 28 以上となる．

■　演習 9　演習 7 を正規近似で求めよ．　■

(4)　正規近似（p値を用いる仮説検定）

棄却域の設定の場合と同様に，p値を用いる仮説検定でも正規近似を用いることもできる．

表示2.15の右の3列にはその結果が求められている．

下側の片側検定のp値は，標準正規分布で$(x + 0.5 - n\pi)/\sqrt{n\pi(1-\pi)}$ 以下の確率が計算されている．

上側の片側検定のp値は，標準正規分布で$(x - 0.5 - n\pi)/\sqrt{n\pi(1-\pi)}$ 以上の確率が計算されている．

両側検定のｐ値は，下側と上側のｐ値の小さい方を2倍する．

■　演習10　　演習8の問題を正規近似で求めよ．　■

(5)　正規近似（連続修正をしない方法）

前の2項では，連続修正を施した正規近似によって棄却域を決め，p値を求めた．

最近の研究で，連続修正を施すことにより，正確法に近い検定結果が得られ，第1種の誤りの確率αを保証することができるが，平均的なαが小さく，保守的な傾向を持つことが問題となっている．その詳細は§2.5(2) で取り上げる．

このような保守的傾向を除くためには，連続修正をしない正規近似を用いるのが好ましいという説がある．

具体的な計算方法は，表示2.15，2.16 のExcelの式から $+0.5$，-0.5 を除けば良い．

(6)　例

コインを100回投げたとき，表の出た回数が39回であった．このコインは歪んでいる（表の出る確率が 50% ではない）といえるであろうか？

帰無仮説 H_0： $\pi = 0.5$，対立仮説 H_1： $\pi \neq 0.5$ として，両側検定をすれば良い．

表示2.15の4行目に示すように，π，n，x に 0.5, 100, 39 を入力する．

両側のｐ値は， 0.0352, 0.0357 で，いずれの方法でも 0.05 以下である．したがって，表の出た割合 0.39 は 0.50 よりも，有意に小さいといえる．

■　演習11　　従来の番組の視聴率は 17% であった．視聴率を上げるためにPR活動を実施した．実施後の視聴率を知るために，ランダムに選んだ500人について調査をした結果，視聴者は 96人，すなわち，視聴率は 19.2% であった．この調査の結果から，視聴率は上がったといえるか？

もし，視聴率に有意差が認められないとしたら，500人中何人以上が視聴したときに有意差ありという結論が出されるか？　■

本日のまとめ

今日学んだ，2 項分布の仮説検定はこの単元の中心課題の一つである．

ここで取り上げた複数の方法の間の関係と，どれを使ったら良いかがはっきりせず，戸惑った人がいたであろう．この点については明日改めて取り上げる．

表示2.15 の使い方を完全にマスターしなければならない．演習以外の問題を自分の周囲から探していろいろ試みることが望まれる．

2.5　仮説検定の方法の比較

(1)　方法の整理

前節では，仮説検定についていくつかの方法を説明した．

それを整理すると，棄却域を決める方法と，p値を計算する方法の2つに，大きく分けられる．それぞれが，2 項分布の正確な確率を計算する方法と，正規分布で近似する方法に分けられる．正規分布による近似では，連続修正を加える方法と省略する方法がある．

棄却域を決める	2 項分布で確率を計算する	
	正規分布で近似する	連続修正を加える
		連続修正を加えない
p値を計算する	2 項分布で確率を計算する	
	正規分布で近似する	連続修正を加える
		連続修正を加えない

棄却域を決める方法は従来標準的な方法として用いられたものである．現在ではp値を計算する方法が主流である．しかし，次節で取り上げる区間推定では，正規近似を使って棄却限界値を計算する式が用いられ，式(2.9) が引用される．

(2)　第1種の誤りの確率

2項分布は観測される値 x が整数に限られる（不連続に変化する）ために，第1種の誤りの確率 α について面倒な問題が生じる．通常の解析では大きな問題

とはならないが，理論的にはいろいろの議論がある．

この項では，この問題を取り上げて，できるだけ分かり易く説明するように努めるが，初学者にとっては難解な点があるであろう．そのような人は，途中の説明を省略し，最後の結論だけを読んでほしい．

前項までは第 1 種の誤りの確率 α を 0.05 を標準的な値として考えてきた．

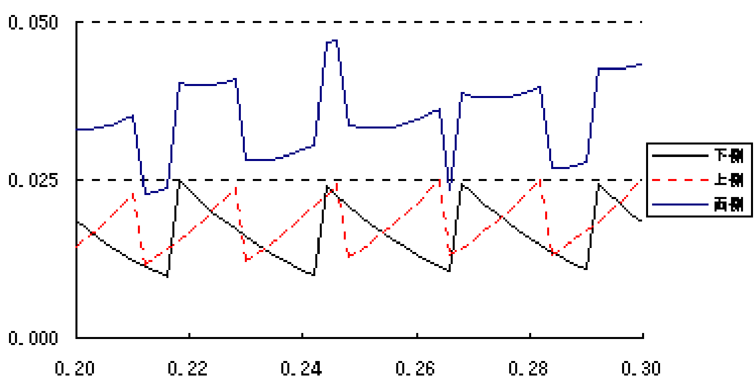
前項で説明した方法は，α が 5% 以下であることを保証する方法である．

$n = 50$ で π を 0.2 から 0.3 まで細かく変化して，第 1 種の誤りの確率を計算して見た．

例えば，$\pi = 0.20$ のとき，棄却域は 4 以下と 17 以上である．帰無仮説が正しいとき，下側または上側に棄却される確率は それぞれ，0.018, 0.014 で，合計 0.033 である．期待される第 1 種の誤りの確率 0.05 に比べてかなり小さい．

横軸に π を，縦軸に 下側で有意となる確率，上側で有意となる確率および両者の合計を取ったグラフを表示2.17 に示す．

表示2.17: 実質的な危険率（正確法）



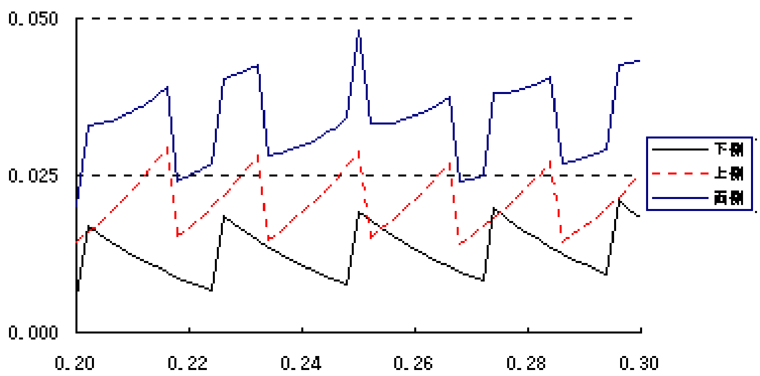
表示2.17 から，片側の危険率は 2.5% を超えることはない，両側の危険率は 5.0% を絶対超えず，平均が 0.35 くらいである．

これから，2 項分布の正確な計算で検定するとき，第 1 種の誤りの確率がα を絶対超えない保証があるが，平均的な危険率は α よりも小さくなることが分か

る．すなわち，保守的な偏りを持つ．

同様の検討を，正規近似で連続修正を加えた場合について試みた結果を表示2.18に示す．

表示2.18: 実質的な危険率（正規近似，連続修正あり）



片側の危険率が鋸状に変化しており，上端は上側は 2.5% を超え，下側は 2.5% よりも下にある．これは，非対称の2項分布を対称の正規分布で近似したことによる偏りである．両側の危険率も不規則に変化しているが 5% は超えない．

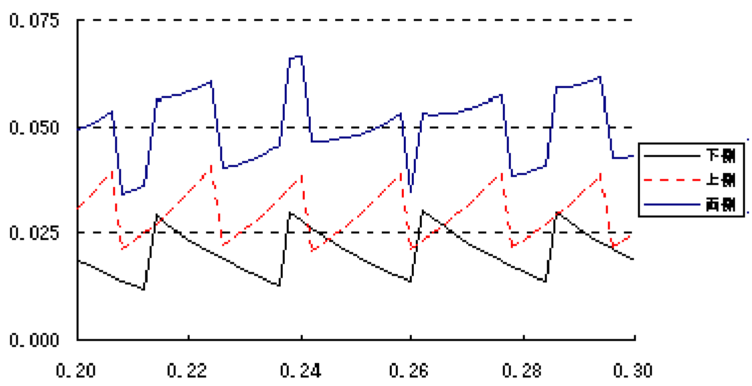
これから，連続修正を加えた正規近似は，両側の危険率は，2項分布による正確な計算と同様，期待される α は保証され，保守的な検定であることが分かる．

連続修正を省略した正規近似について同様の計算をすると，表示2.19 が得られる．

表示2.19を前の２つのグラフと比較すると，鋸状の折れ線は，片側の危険率は 2.5%，両側の危険率は 5.0% の前後で変化している．両側の危険率の平均はほぼ 5.0% である．

最後のグラフを見ると，上側と下側で平均的に 2.5% から上下に外れている．この偏りを修正する方法に mid-P 法という方法が提案されているが，ここでは説明を省略する．

表示2.19: 実質的な危険率（正規近似，連続修正なし）



以上の3つのグラフによって仮説検定の特徴を見ることができる．

それでは，現実の問題にはどの方法を使ったら良いであろうか？

新しい医薬品の開発過程で，結果を行政当局に提出して認可を取るというように第1種の誤りの確率を厳格に保証しなければならない場合は，2項分布の正確な計算を用いる必要があるであろう．

それに対して，企業内の決定のための結論を出す場合のように，平均的な危険率を保証すれば良い場合には，連続修正を省略した正規近似で十分であると思われる．

なお，ここに述べた判断については，業界による習慣的な，または，協定された方法の有無を調査して，適切な方法を選択するのが良いであろう．

本日のまとめ

今日の前半では，仮説検定の方法の違いを明らかにした．この違いは理解してほしい．

後半で取り上げた実質的な危険率については，数理統計学者の間でも議論の残るところである．最後に記したように，受講生の置かれている環境に応じて適切な判断をしてほしい．

「第1種の誤りの確率」の項を読み飛ばした人，十分に理解できなかった人は，実務の中で割合の検定や推定の経験を積んだ後に読み直すことにより，役に立つであろう．

3　2項分布に関する推定と検定 (2)

3.1　割合の区間推定 (1)

(1)　簡単な例

§1.1 で，100 例の手術で全例成功し，実績での失敗率が 0% であるとき,「本当の失敗率は 0.03 未満であろう」という結論を導いた.

上の 0.03 は，次のように考えて求めた.

真の失敗率が $\pi = 0.03$ であるとき，$n = 100$ のサンプルで失敗数が $x = 0$ となる確率は，0.048 で，この値は十分に小さいから，$\pi = 0.03$ 以上である可能性は低い.

これを，仮説検定と結びつける.

$n = 100,\ x = 0$ のデータが得られたとき，帰無仮説 $\mathrm{H}_0:\ \pi = 0.03$ を検定する.

帰無仮説が正しいとき，$x \leq 0$ となる確率は $(1-0.03)^{100} = 0.048$ で，この値（p値）は 0.05 よりも小さいので，帰無仮説は棄却される（ただし，片側検定とする）.

ここで，$\pi = 0.02$ とすると，$x = 0$ となる確率は 0.05 よりも大きくなり，仮説は棄却されない.

これから，「真の割合 π は，仮説検定で棄却されない範囲に含まれる」と考えれば，ほぼ 間違いないと思われる．ここで，ほぼ という仮説検定でも用いられたあいまいな言葉を使った．これを，仮説検定の第 1 種の誤りの確率 α と対応させると，1 からほぼの程度 (5%) を引いた 0.95 の確かさで，$\pi < 0.03$ であるといえる．この 0.95 を 信頼率 という.

上に示した $\pi < 0.03$ は概数である．もっと正確な値は，$x = 0$ となる確率がちょうど 0.05 になる π を求めれば良い.

これは，

$$(1-\pi)^{100} = 0.050$$

から導かれる，

$$\begin{aligned}100\log(1-\pi) &= \log(0.050) = -1.3010\\ 1-\pi &= 10^{-0.01301} = 0.97049\end{aligned}$$

から，$\pi = 0.0295$ が得られる．

これから，$\pi < 0.0295$ という範囲が求められる．このような推定を 区間推定 と呼び，この範囲を 信頼区間 という．また，区間の境界値を 信頼限界 と呼ぶことがある．

上の例では π が <u>0.03 以下である</u> という信頼区間を求めた．このように，上限値以下，または，下限値以上という場合を片側区間推定 と呼ぶ．

下限値以上で かつ 上限値以下 という信頼区間を求める場合を両側区間推定と呼ぶ．

目的に応じて両者を正しく使い分けることが大切である．

この方法を一般の x の場合に適用するためには，計算方法に工夫が必要である．その具体的な方法は§3.2 (3) で取り上げる．

(2)　正規近似（単純法）

前章の最後で，実用的には連続修正を加えない正規近似が良いであろうと述べた．そこで，この節から，次の §3.2(1) まで，この方法による区間推定の方法を説明し，それ以外の方法は，§3.2(2) 以降に補足的に説明する．

正規近似の仮説検定の式(2.9) (p.46)，

$$x \sim n\pi \pm (1.96\sqrt{n\pi(1-\pi)} + 0.5)$$

から，連続修正の ± 0.5 項を除き，不等式の両辺を n で割ると

$$\frac{x}{n} = p \sim \pi \pm 1.96\sqrt{\frac{\pi(1-\pi)}{n}} \tag{3.1}$$

が得られる．

この式は，帰無仮説（母集団の割合が π である）が正しいとき，サンプルの割合 p の95%が含まれる範囲を与えるものである．

これを，π が左辺に来るように変形すると，

$$\pi \sim p \pm 1.96\sqrt{\frac{\pi(1-\pi)}{n}} \tag{3.2}$$

が得られる．

この式は，サンプルの割合 p が得られたとき，母集団の割合 π の95%が含まれる範囲を与えるものである．

これから π の信頼区間を求めようとすると，両辺に π が含まれるので，面倒である．そこで，右辺の π を p に置き換えた次の式がこれまで広く用いられていた．

$$\pi \sim p \pm 1.96\sqrt{\frac{p(1-p)}{n}} \tag{3.3}$$

$n = 100,\ x = 30,\ p = 30/100 = 0.3$ のとき，

$$\pi \sim 0.3 \pm 1.96\sqrt{\frac{0.3(1-0.3)}{100}} = 0.3 \pm 0.090 = (0.210,\ 0.390)$$

となる．

表示3.1: 正規近似による区間推定

n	x	p	片側α	下限	上限	
100	30	0.30	0.025	0.210	0.390	1.960
100	0	0.00	0.025	0.000	0.000	1.960
100	1	0.01	0.025	-0.010	0.030	1.960

ここで，$n = 100,\ x = 0,\ p = 0$ の場合を考えると，

$$\pi \sim 0 \pm 0, \quad (0 \le \pi \le 0)$$

という結果が得られる．

また，$x = 1$ とすると，

$$\pi \sim 0.01 \pm 0.020, \quad (-0.010 \le \pi \le 0.030)$$

となり，信頼区間の範囲に負の値が含まれる．

これらの結果は不自然であり，連続修正を加えても解決されない．

(3) 正規近似（スコア法）

式(3.2) を π について直接解く方法が考えられる．

式(3.2) 右辺の p を左辺に移項し，両辺を2乗して整理すると，2次方程式が導かれる．

$$\pi - p = \pm 1.96\sqrt{\frac{\pi(1-\pi)}{n}}$$
$$n(\pi - p)^2 = 1.96^2\pi(1-\pi)$$
$$(n + 1.96^2)\pi^2 - 2(np + 1.96^2/2)\pi + np^2 = 0$$
$$(n + 1.96^2)\pi^2 - 2(x + 1.96^2/2)\pi + x^2/n = 0$$

$n = 100$ で，$x = 30$ とすると，

$$\begin{array}{rrrr}
(100+1.96^2)\pi^2 & -2(30+1.96^2/2)\pi & +30^2/100 & = 0 \\
a \quad \pi^2 & -2 \quad b \quad \pi & + \quad c & = 0 \\
103.841\pi^2 & -2 * 31.921\pi & +9.000 & = 0
\end{array}$$

となり，その解は

$$\pi = \frac{b \pm \sqrt{b^2 - ac}}{a} = (0.219,\ 0.396)$$

となる．

表示3.2: スコア法による区間推定

n	x	p	片側α	下限	上限		a	b	c	√(b^2-ac)
100	30	0.30	0.025	0.219	0.396	1.960	103.841	31.921	9.000	9.185
100	0	0.00	0.025	0.000	0.037	1.960	103.841	1.921	0.000	1.921
100	1	0.01	0.025	0.002	0.054	1.960	103.841	2.921	0.010	2.737

この結果を，前の方法の結果 0.210, 0.390 と比較すると，前の方法は，p の両側に等間隔の幅をつけているのに対して，この方法は，両側の幅が異なる．

ここで，$n = 100$ で，$x = 0$ とすると，

$$(100 + 1.96^2)\pi^2 - 2(0 + 1.96^2/2)\pi + 0^2/100 = 0$$
$$103.841\pi^2 - 2 * 1.921\pi + 0.00 = 0$$

となり，その解は

$$\pi = 2 * 1.921/103.841 = 0.0370$$

すなわち，$\pi < 0.0370$ となる．

$x = 1$ とすると，その解は

$$\pi = (0.002,\ 0.054)$$

となる．

この方法によって，式 (3.3) を用いる方法の不都合が解消される．

この方法は スコア法 と呼ばれる．表示 3.2 には式に忠実に従って計算する方法が示されている．もっと実用的な具体的な計算方法は §3.2 で詳しく説明する．

(4)　信頼区間の概数

スコア法による 95% 信頼区間がどのくらいの幅になるのかをグラフで示したのが，表示 3.3 である．

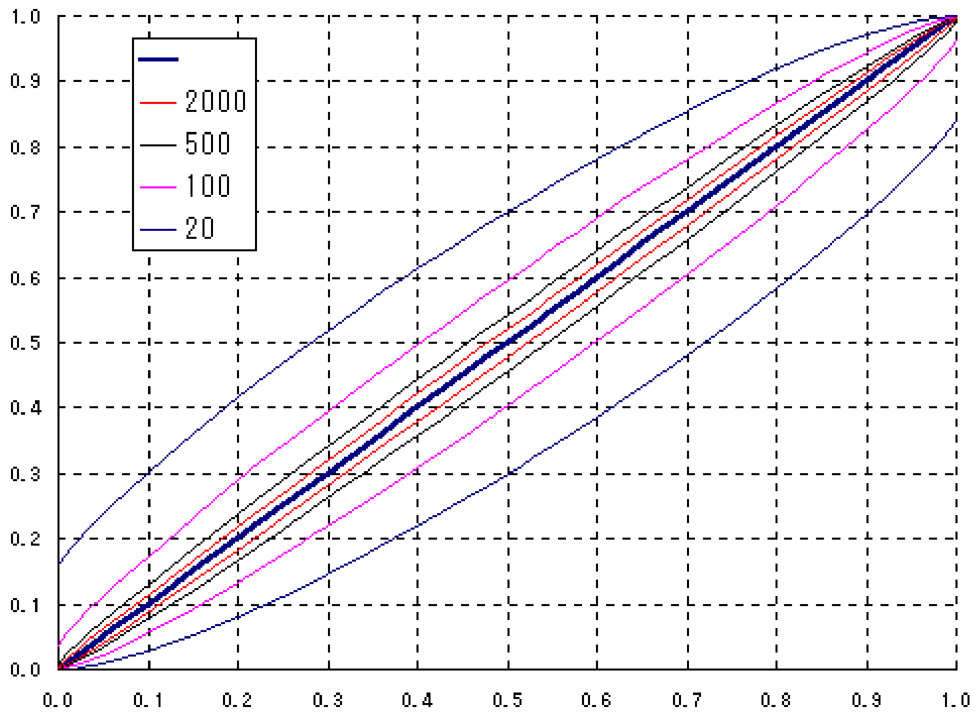
横軸に観測された割合 $p = x/n$ が，縦軸に区間推定の上限と下限が取られている．

$n = 20,\ 100,\ 500,\ 2000$ の 4 対の曲線が描かれている[1]．

$n = 20,\ x = 6,\ p = 6/20 = 0.3$ が得られたとき，π の信頼区間は，横軸が 0.3 で，一番上と一番下の曲線の値を読み取る．0.52, 0.15 が得られる．区間の

[1] 信頼率や n を変えたグラフを見たいときは，Excel 画面のグラフの左の計算表のパラメータを修正する．

表示3.3: スコア法による信頼区間



上限と下限の比は $0.52/0.15 = 3.6$ であり，このような区間推定は広すぎて実用にならないであろう．

それでは，$n = 100$，$x = 30$，$p = 30/100 = 0.3$ に対する信頼区間はいくらになるかを見よう．外側から2番目の曲線の値を読むと，$0.22 \le \pi \le 0.40$ である．上限と下限の比は $0.40/0.22 = 1.81$ となる．大まかにいうと,「真の割合は 20% から 40% の範囲内である」という不確かな結果しか得られない．この幅は皆さんの予想と合っているであろうか？　恐らく予想以上に広いと感じた人が多いであろう．

もっと大きな n について調べて見よう．$n = 2000$，$x = 600$，$p = 600/2000 = 0.3$ の場合は，$0.28 \le \pi \le 0.32$ となる．

このグラフを利用する別の例を取り上げる．π が 20% 前後と予想される．π

の信頼区間の範囲が ± 0.05 程度の精度で推定したい．どのくらいの n が必要だろうか？

表示3.3 で横軸が 0.2 の縦軸を見ると，0.15, 0.25 は，$n = 100$，500 の線の間である．したがって，n は 100〜500 の範囲内となる．もう少し精度を高めて n を求めたい．

表示3.3 の Excel シートで，グラフの左の計算表の表頭の n を変更すると，グラフも自動的に変化する．n を 300 に変更すると，2本の曲線が 0.15, 0.25 の近くを通る．これから，必要な n は300程度と予想される．

表示3.3 から，このように区間推定の精度の概数を知ることができる．

本日のまとめ

今日は，比率の区間推定の基本的な方法を学んだ．具体的な計算方法は明日詳しく学ぶが，今日は基本的な考え方を理解してほしい．

最後の項で，区間推定の幅がどのくらい広いかを見ることができるグラフを示した．このグラフのイメージを頭に入れておくことにより，自分のデータがどのくらいの信頼性を持っているか，期待する精度で推定するためにはどのくらいの n が必要かを知ることができるであろう．

3.2　割合の区間推定 (2)

(1)　ゴールシークによるスコア法の計算

スコア法で区間推定をするために2次方程式を解くのは面倒である．方程式を解かなくても解が求められる便利な方法を説明する．

§3.1(2) では，仮説検定で棄却域を求める式 (3.1) を変化させて信頼区間を求

める式(3.2)を導いた.

仮説検定と区間推定は，このように 表裏の関係にある.

p 値による仮説検定を裏から眺めることにする.

仮説検定で p 値を計算する表示2.15(p.45) から正規近似の下側と上側だけをとりだし，連続修正を除くと表示3.4 が得られる.

表示3.4: スコア法による区間推定

	A	B	C	D	E	F	G
3	π	n	x	下側	上側		
4	0.2	100	11	0.0122	0.9878		
5	0.2	100	12	0.0228	0.9772		
6	0.25	100	30	0.8759	0.1241		
7	0.35	100	30	0.1473	0.8527		
8	0.2190	100	30	0.9750	0.0250		
9	0.3958	100	30	0.0250	0.9750		
10							
11	D4:	=NORMSDIST((C4-B4*A4)/SQRT(B4*A4*(1-A4)))					
12	E4:	=1-NORMSDIST((C4-B4*A4)/SQRT(B4*A4*(1-A4)))					

上の2行は表示2.15 と π, n, x は同じであるが，連続修正を除いているので，値が少し異なる.

中央の2行の $n = 100$, $x = 30$ は 表示3.2 の1行目と同じである.

π の列には適当な値（ここでは， 0.25, 0.35）を入力する.

ここで，この π の値を変化させると下側と上側の p 値が変化する．試行錯誤で 上側 または 下側 の p 値が 0.025 に近くなる π を求める.

下の2行はいずれかの p 値がちょうど 0.025 になっている．ここに得られた π は表示3.2 の結果と一致している.

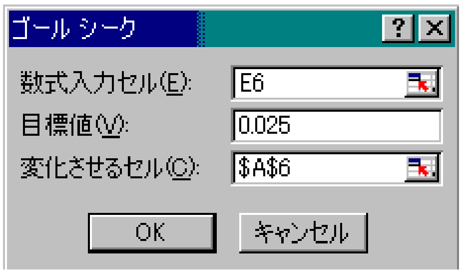
これだけの精度の π を試行錯誤で求めるのは，大変な労力を要し，現実には不可能である．この試行錯誤を自動的に実行してくれるのが，Excel の ゴールシーク である.

まず，π のセルに初期値（例えば，$p = x/n = 0.30$）を入力する．0.025 にしたいセル（例えば，E6）をクリックしてから，トップメニューから「ツール」

「ゴールシーク」を選択する．

表示3.5 の画面が表示される．

表示3.5: ゴールシーク設定画面



「数式入力セル」には最初に指定したセル（E6）が表示されているはずである（違う場合は，修正することができる）.「目標値」に 0.025 を入力する.「変化させるセル」には，π のセル（ここでは A6）を入力する．

「OK」をクリックする.「解答がみつかりました」と表示されたら,「OK」をクリックする．

このような手順で求められたのが，表示3.4 の下の 2 行で $0.2190 \leq \pi \leq 0.3958$ となる[2]．

ゴールシークは，スコア法だけでなく，次項の正確法でも用いられる．

(2)　正確法

§2.5 の最後で，仮説検定の方法がいくつかあるが，一般には連続修正を加えない正規近似が良いであろうと述べ，場合によっては，2 項分布の正確な確率を使う必要があることを注意した．

このような特別な場合のために，正確な方法を説明する．

---

[2] ゴールシークは逐次近似法を用いているので，いつも同じ解が得られるとは限らない．別の初期値を設定すると異なる解の得られることがある．逐次近似の精度を変更するためには,「ツール」,「オプション」,「計算方法」,「反復計算」で,「変化の最大値」を 0.0001 のように小さく設定する．

スコア法の場合と同様に，表示2.15 (p.45) から 2 項分布 の欄を取り出して表示3.4 と同じ形の表を準備する．

表示3.6: 正確法による区間推定

π	n	x	下側	上側
0.2	100	11	0.0126	0.9943
0.2	100	12	0.0253	0.9874
0.25	100	30	0.8962	0.1495
0.35	100	30	0.1730	0.8764
0.2124	100	30	0.9856	0.0250
0.3998	100	30	0.0250	0.9851

スコア法の場合と全く同じ手順で，ゴールシークを使って，信頼区間を求めることができる．結果は $0.2124 \leq \pi \leq 0.3998$ となる．

■ **演習 12** $n = 100,\ x = 30$ が得られたとき，これから，片側の区間推定を求めたい．信頼率が95% の 下側 または 上側 の信頼区間を，スコア法と正確法で計算せよ． ■

(3) VBA マクロを用いる計算表

これまでに説明したように，ゴールシークを使ったり，2 次方程式を解いて，信頼区間を求めるのは煩雑である．

そこで，$n,\ x$ と危険率 α（片側）を入力すると，信頼区間が求められる計算表を準備した．

この計算表では，Excel の VBA で書かれた 2 つのユーザー定義関数

=Bin_Ac(x, n, α, "L" または "U") 正確法

=Bin_Sc(x, n, α, "L" または "U") スコア法

が用いられている．

最後のパラメータを "L" とすると下側の，"U" とすると上側の信頼限界が得られる（L, U の左右を " " で囲み，文字として指定しなければならない）．

表示3.4，表示3.6 と同じ結果が得られている．

表示3.7: VBA マクロによる信頼区間の計算表

	A	B	C	D	E	F	G	H	I	J
28				スコア法		正確法		上/下		
29	n	x	α(片側)	下側	上側	下側	上側	スコア法	正確法	正/スコア
30	100	30	0.025	0.2189	0.3958	0.2124	0.3998	1.808	1.882	1.04
31	100	30	0.05	0.2307	0.3798	0.2249	0.3842	1.646	1.708	1.04
32	100	10	0.025	0.0552	0.1744	0.0490	0.1762	3.157	3.596	1.14
33	100	5	0.025	0.0215	0.1118	0.0164	0.1128	5.187	6.867	1.32
34	20	1	0.025	0.0089	0.2361	0.0013	0.2487	26.587	196.613	7.40
35	2371	1040	0.025	0.4188	0.4587	0.4185	0.4589	1.095	1.096	1.00
36										
37	D30:	=Bin_Sc($B30,$A30,$C30,"L")								
38	E30:	=Bin_Sc($B30,$A30,$C30,"U")								
39	F30:	=Bin_Ac($B30,$A30,$C30,"L")								
40	G30:	=Bin_Ac($B30,$A30,$C30,"U")								

信頼率が95%の両側信頼区間の場合には，α (片側) の欄に0.025 を入力する．

VBA マクロを使うための注意

提供するExcelファイル「第3単元.XLS」には，マクロプログラムが記録される．このファイルの中では，新しいシートを作っても，これらのマクロを使うことができる．新しいファイル（ブック）を作って，その中で使おうとすると，結果が得られない．その対策については，テキスト冒頭の「学習の手引き」を参照のこと．

(4)　スコア法と正確法の比較

説明に用いた $n = 100,\ x = 30,\ p = 0.3$ の場合について，スコア法と正確法の信頼区間を比較すると，正確法の方が広くなっている．§2.5(2) で説明したように，正確法は信頼率が常に $1 - \alpha$ 以上であることが保証されているのに対して，スコア法は信頼率が平均的に $1 - \alpha$ になるためである．

この例では，両者の違いがそれほど大きくない．

両者の違いはどのような場合に，どの程度生じるかを調べるために，表示3.7

の右には，スコア法と正確法の結果を比較するために，信頼区間の 上限/下限の比が計算されている．

$n = 100$ で，$x = 10, p = 0.10$，$x = 5, p = 0.05$ と p を小さくすると，上下の比が 14%，32% 違ってくる．さらに，$n = 20, x = 1, p = 0.05$ とすると，比が 7.4 倍という大きな違いが見られる．しかし，これらの場合は，信頼区間が，$5 \sim 18\%$，$2 \sim 11\%$，$0.1 \sim 25\%$ と極めて広いので，実用的には，近似であるスコア法で十分であろう．

(5) 区間推定の事例

新聞の世論調査の結果として,「内閣支持率が 43.86% であった」と書かれていた．

この調査は 2371 人の回答者から得られたもので，1040 人が内閣を支持した．この値を表示 3.7 に入力すると，最後の行に，スコア法の信頼区間として $41.88\% < \pi < 45.87\%$ が得られる．

すなわち，真の支持率は 41.88% 〜 45.87% という広い範囲のどこかである．このような場合に 43.86% というように報告することは，結果の精度について誤解を与える危険がある．前にも述べたように，44% というように，無意味な小数点以下の数値を省略することが望ましい．

本日のまとめ

昨日は区間推定の基本的な考え方を説明した．そこで,「こんな面倒な計算はとてもできない」と感じた方もおられたであろう．

今日は，ゴールシークという強力な武器を使うと，かなり簡単に信頼区間が得られることを知った．さらに，VBA で書かれたマクロプログラムを使うと，もっと簡単になることを知り，これなら使えると感じられたことであろう．

VBA マクロを自分のプログラムで自由に使えるようにするためには，Excel の操作にある程度習熟する必要がある．

3.3　検出力とサンプルサイズの計算

(1)　賭けの不正を見抜くには

コインの裏表の賭けで，勝率が 40% になるようにコインに細工をしたときに見破れるであろうか？

$n = 10$ の勝数 x について，帰無仮説 $\mathrm{H}_0:\ \pi = 0.5$ を 対立仮説 $\mathrm{H}_1:\ \pi < 0.5$ に対して片側検定するとき，棄却域は $x \leq 2$ で，第1種の誤りの確率は 0.055 であることは，§1.4(1) で学んだ（この確率は通常の $\alpha = 0.05$ より少し大きいが，この棄却域を用いることにする）．

それでは，もし，$\pi = 0.4$ となるような不正をされたとき，帰無仮説が棄却される，すなわち，不正を見破る確率はいくらになるであろうか．

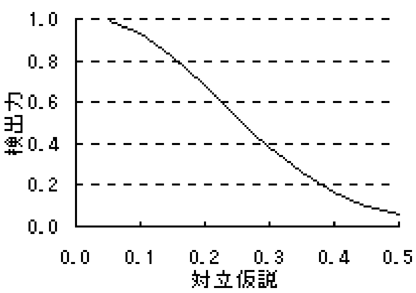
$\pi = 0.4,\ n = 10$ で $x \leq 2$ となる確率は 0.167 である．

すなわち，$n = 10$ の観測では，この細工は 0.167 の割合でしか発見できない．逆にいうと，$1 - 0.167 = 0.833$ の割合で見逃される．この値を **第2種の誤りの確率** と呼び，β で表わすことはすでに §1.4 (2) で述べた．また，$1 - \beta$ の 0.1672 は **検出力** と呼ばれる．

表示3.8: π と検出力の関係

x	n	π	検出力	β
2	10	0.50	**0.055**	0.945
2	10	0.45	0.100	0.900
2	10	0.40	0.167	0.833
2	10	0.35	0.262	0.738
2	10	0.30	0.383	0.617
2	10	0.25	0.526	0.474
2	10	0.20	0.678	0.322
2	10	0.15	0.820	0.180
2	10	0.10	0.930	0.070
2	10	0.05	0.988	0.012

太字は第1種の誤りα



$\pi = 0.3$ になると，$x \leq 2$ となる確率，すなわち 発見できる確率は 0.3828 に増加するが，見逃しの確率 β はまだ 0.6172 で，半分以上である．

π を小さくしたときに検出力がどのように変化するかを計算し，グラフ化し

たのが表示3.8である．

π が 0.15 まで小さくなると，検出力が 80% 以上となる．

このグラフは **検出力曲線** と呼ばれ，検定の特性を表わすために広く用いられる．

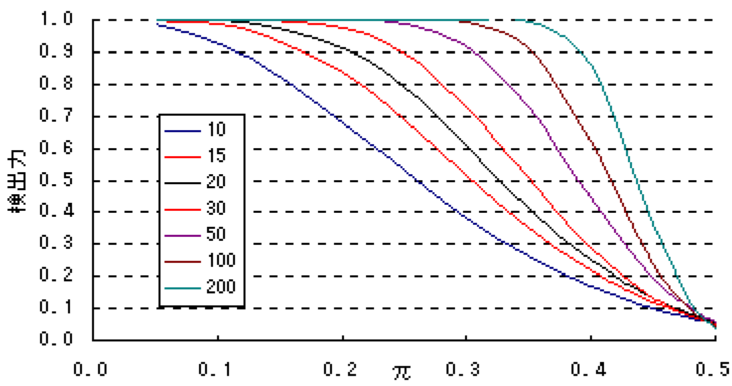
$n = 10$ の観測では極端に細工をした場合でないと発見できないことが分かった．

細工を見抜くためには，n をもっと大きくしなければならない．

n を 10, 15, ..., 200 と大幅に変化させて，それぞれに対する検出力と，検出力曲線を表示3.9に示す．

表示3.9: 検出力と検出力曲線

		π									
n	x	0.50	0.48	0.45	0.40	0.35	0.30	0.25	0.20	0.12	0.05
10	2	**0.055**	0.070	0.100	**0.167**	0.262	**0.383**	**0.526**	0.678	0.891	0.988
15	4	0.059	0.080	0.120	0.217	0.352	0.515	0.686	0.836	0.974	0.999
20	6	0.058	0.081	0.130	0.250	0.417	0.608	0.786	0.913	0.993	1.000
30	10	0.049	0.076	0.135	0.291	0.508	0.730	0.894	0.974	1.000	1.000
50	19	0.059	0.101	0.197	0.446	0.726	**0.915**	0.986	0.999	1.000	1.000
100	41	0.044	0.096	0.241	0.623	0.912	0.993	1.000	1.000	1.000	1.000
200	87	0.038	0.114	0.362	**0.860**	0.995	1.000	1.000	1.000	1.000	1.000



ここで，有意となる x の上限値は，α（$\pi = 0.5$ に対する検出力）が 0.05 に近い値になるように選んだ．

表示3.9から，$\pi = 0.4$ を 90% の確率で検出するためには，n が 200 でも足りないことが分かる．

(2)　正規近似による検出力の計算

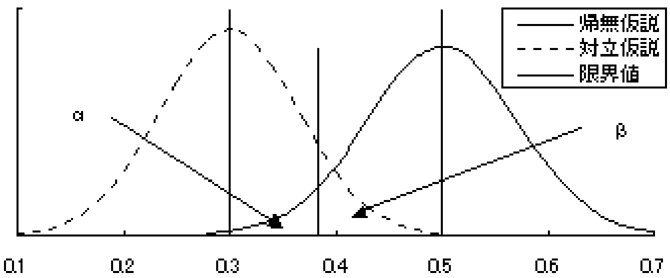
n がある程度大きい場合には，2項分布を正規分布で近似して，検出力の近似値を求めれば，実用的に十分である．

例として，$\mathrm{H}_0:\ \pi = 0.5 \equiv \pi_0,\ \alpha = 0.05,\ n = 50$ の下側片側検定で，$\mathrm{H}_1:\ \pi = 0.3 \equiv \pi_1$ の検出力を求めてみる[3]．以下，帰無仮説の π と対立仮説の π を区別するために，π_0，π_1 で表わす．

帰無仮説と対立仮説の下での p の分布を正規分布で近似する(連続修正は省略)と，表示3.10となる．

表示3.10: 正規近似による検出力

n	π0	π1	α(片側)	限界値	β
50	0.5	0.3	0.05	0.384	0.098
	0.0707	0.0648			



それぞれの分布の中心に縦線が引かれている．対立仮説(左の点線，$\pi_1 = 0.3$)の方が標準偏差が小さいので，分布の山が高くなっている．

棄却域は，帰無仮説（右の実線）の分布で，下側の確率が α となる領域である．棄却限界値に縦線が引かれている．この縦線の左の面積が α である．限界値の p は

$$p = \pi_0 - 1.645\sqrt{\frac{\pi_0(1-\pi_0)}{n}} = 0.5 - 0.116 = 0.384$$

で計算される．

---

[3] $\pi = 0.5 \equiv \pi_0$ の $\equiv$ は，「0.5 を一般化して π_0 とする」ことを表わす記号である．

対立仮説（左の点線）の分布で，限界線の右側の面積は，帰無仮説正しくないにもかかわらず棄却されない確率，すなわち，第2種の誤りの確率 β である．

β を求める式は次のように考えて導かれる．

限界値 と π_1 との距離は，対立仮説の下での標準偏差の何倍かを求める．

$$z = \frac{\text{限界値} - \pi_1}{\sqrt{\pi_1(1-\pi_1)/n}} = \frac{0.384 - 0.3}{\sqrt{0.3(1-0.3)/50}} = 1.291$$

標準正規分布で，$z = 1.291$ よりも大きい確率を計算すると，0.098 となる．これが，第2種の誤りの確率 β で，検出力は $1 - 0.098 = 0.902$ となる．

この値は，表示3.9 の 0.915 に対応する．

表示3.10 で，グラフの上の n，π_0，π_1，α を修正すると，その右の値とグラフが自動的に更新される．これを使って，これらのパラメータと検出力の関係を視覚によって理解することができるであろう．

■ **演習 13** 表示3.10 で，n，π_1 を変化させて，グラフの変化を眺めよ．また，得られた検出力の値を表示3.9 の2項分布で計算した検出力の値と比較し，正規分布近似の近似の程度を確認せよ． ■

これまでは，$\pi_1 < \pi_0$ の場合について説明した．$\pi_1 > \pi_0$ の場合にも検出力 β が求められるように，拡張した検出力を求める計算表を表示3.11 に示す．

表示3.11: 正規近似による検出力の計算表

	AB	AC	AD	AE	AF
3	n	π0	π1	α(片側)	β
4	50	0.5	0.3	0.05	0.098
5	50	0.5	0.7	0.05	0.098
6	50	0.5	0.3	0.05	0.098
7	60	0.5	0.3	0.05	0.056
8	70	0.5	0.3	0.05	0.032
9	61	0.5	0.3	0.05	0.053
10	62	0.5	0.3	0.05	0.050
11	62.11	0.5	0.3	0.05	0.050
12	100	0.2	0.25	0.1	0.512
13	200	0.2	0.25	0.1	0.327
14	307.63	0.2	0.25	0.1	0.200

β の求められているセル AF4 には

```
=NORMSDIST(SIGN(AC4-AD4)*(AD4-(AC4-SIGN(AC4-AD4)
    *NORMSINV(1-AE4)
    *SQRT(AC4*(1-AC4)/AB4)))/SQRT(AD4*(1-AD4)/AB4))
```

が入力されている．

SIGN(AC4-AD4) は $\pi_0 - \pi_1$ の符号により，$+1$ または -1 とする関数である．

$\pi_1 = 0.7$ としても同じ検出力が得られる．

(3)　サンプルサイズの求め方

上記の検討から，$n = 50$ では，π の変化を見逃す確率 β が 0.098 であることが分かった．

β を 0.05 と小さくするためには，n をどれだけ増やせば良いであろうか．

表示3.10 の計算表は，左の4項目（n, π_0, π_1, α(片側)）を修正すると，β が再計算されるようになっている．

そこで，n の値 50 を 60, 70 と増やしてみる．$n = 70$ とすると，$\beta = 0.032$ と 0.05 以下になる．これから，必要な n は 60 と 70 の間にあることが分かる．

今度は n を 1 ずつ修正すると $n = 62$ のときに，$\beta = 0.050$ となることが分かる．

このような試行錯誤によって，第2種の誤りの確率 β を希望する値以下に抑える（逆にいうと，検出率を $1-\beta$ にする）ために必要な n （サンプルサイズ）を求めることができる．

§3.2(1) で用いた ゴールシーク を使えば，もっと簡単にサンプルサイズを決めることができる．ただし，$n = 62.11$ と小数点以下の端数が付くので，切上げて 63 とする．

(4)　適用例

あるテレビ局で，視聴率が 20% を超えたら 金一封 が配られるものとする（仮想例）．

100人の視聴者についての調査した結果で，判断するものとすると，いろいろの基準が考えられる．

100人の 20% を超える，すなわち，$x > 20$ のとき金一封を配るとすると，真の視聴率 π が 0.20 でも $x > 20$ の確率は2項分布の計算で 0.4405 となる．視聴率が 20% を超えていないにもかかわらず，50%近くの確率で報奨金を支払うのは好ましくない．

そこで，π が 0.20 を超えたことが ほぼ 確かであれば，報奨金を払うことにする．そのためには，帰無仮説 $\mathrm{H}_0:\ \pi = 0.20$ を，対立仮説 $\mathrm{H}_1:\ \pi > 0.20$ に対して検定して，有意であるという結論が得られたときに，報奨金を払うことにすれば良い．ここで ほぼ として 90%（$\alpha = 0.10$）と通常の検定よりもゆるく考える．帰無仮説のもとで，$x \geq 26$ の確率が 0.0875 と 0.10 以下になるので，100人中26人以上が視聴しているときに報奨金が払われることになる．

支払う側からすれば，無駄に報奨金を支払う危険は減少するが，受け取る側から考えると，真の視聴率 π が 20% よりかなり大きくならないと報奨金がもらえなくなる．

真の視聴率が 25% になったら ほぼ 報奨金が貰えるようなシステムを作ってほしいという希望が出された．ここで，ほぼ の程度として 80% を考える．

そのためには，視聴率調査のサンプルサイズ n を増やす必要がある．それでは，n をどのくらいまで増やさなければならないであろうか．

表示3.11で，$n = 100,\ \pi_0 = 0.2,\ \pi_1 = 0.25,\ \alpha = 0.1$ を入力すると，$\beta = 0.512$ と $n = 100$ では報奨金を受け取れない確率 β が半分以上であることが分かる．

そこで，n を大きくして，$\beta = 0.20$ 以下になる n を求める．その結果，$n = 308$ のサンプルについて視聴率を調査することにより，支払う側と受け取る側の両方がほぼ満足する結果の得られることが分かる（表示3.11の下参照）．

■ **演習14** テレビの視聴率 π が 20% を割ったことが確かめられたら，連続番組を打ち切ることにしたい．

また，視聴率 π が 15% を割ったら，ほぼ 確実に，番組を打ち切りたい．

$\alpha,\ \beta$ をいろいろと変化させて，視聴率調査に必要な n を求めよ．■

本日のまとめ

今日は，希望する検出力を確保するためには，サンプルサイズ n をいくらにすれば良いかについての考え方と，そのための計算表を学んだ．ここに示されている演習を解くだけでなく，自分の問題と結びつけて必要な n を計算して見よ．おそらく，従来考えていた以上の n が必要であることが分かるであろう．

3.4　割合の差の推定と検定

(1)　例1

肺がん患者100人と非患者300人について，喫煙本数を調べたところ，1日に40本以上吸う重喫煙者が，患者で30人，非患者で54人あった．患者の方が非患者より重喫煙者の割合が大きいといえるか．

前節では一つの母集団について，割合の推定・検定の方法を学んだが，この節では2つの母集団について，割合の差がいくらであるかの推定や，割合に差があるか否かの検定の方法を考えよう．

患者，非患者の標本で，重喫煙者の割合は，

患者　$p_1 = 30/100 = 0.30$

非患者　$p_2 = 54/300 = 0.18$

全体　$p = 84/400 = 0.21$

となる．これから，両者の割合の差 d は

$$d = p_1 - p_2 = 0.30 - 0.18 = 0.12$$

と推定される．

これは，割合の差の **点推定値** である．それでは，割合の差が統計的に意味があるかどうか，また，割合の差の区間推定はどのようにして導いたら良いであろうか．

この問題を正確に解くことは面倒なので，通常は正規近似が用いられる．

(2)　仮説検定

2 項分布の性質から，2 つの割合 p_1，p_2 の分散は

$$V[p_1] = \frac{\pi_1(1-\pi_1)}{n_1}$$
$$V[p_2] = \frac{\pi_2(1-\pi_2)}{n_2}$$

である．p_1，p_2 は別のサンプルから求めたものであるから，独立であり，分散の加法性により，差 $d = p_1 - p_2$ の分散はそれぞれの分散の和となる．

$$V[d] = \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2} \tag{3.4}$$

となる．

帰無仮説を $H_0:\ \pi_1 = \pi_2 = \pi$ として，検定する．

帰無仮説のもとでの d の分布を考えると，期待値は 0 で，分散は

$$\begin{aligned} V[d] &= \frac{\pi(1-\pi)}{n_1} + \frac{\pi(1-\pi)}{n_2} \\ &= \pi(1-\pi)\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \end{aligned} \tag{3.5}$$

となる．ここで，π の値は分からないので，その推定値 p （全体での患者の割合）で代用することにする．

$$\begin{aligned} V[d] &= p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \\ &= 0.21(1-0.21)\left(\frac{1}{100} + \frac{1}{300}\right) = 0.0022 \end{aligned}$$

この平方根を取ると，$d = p_1 - p_2$ の標準偏差が求められる．d のように，推定や検定に用いられる量の標準偏差をとくに **標準誤差** （Standard Error） と呼ぶことがある．以下，標準誤差を $se()$ で表わす．

$$se(d) = \sqrt{0.0022} = 0.0470$$

このようにして，割合の差 d の標準誤差 $se(d)$ が求められたので，差 d をその標準誤差で割って基準化する．

$$z = \frac{d}{se(d)} = \frac{0.12}{0.0470} = 2.551$$

標準正規分布で ± 2.551 の外に出る確率（両側p値）は 0.011 である．この値は 0.05 より小さいので，$\alpha = 0.05$ で有意差のあることが確かめられた．

以上の計算過程を表示3.12の上 2/3 に示す．

表示3.12: 2つの割合の差の検定と区間推定（例1）

	n	x	p
患者	100	30	0.30
非患者	300	54	0.18
合計	400	84	0.21

仮説検定

d	V[d]	se(d)	z	p値(片側)	p値(両側)
0.12	0.0022	0.0470	2.551	0.005	0.011

区間推定

d	V[d]	se(d)	α(片側)	dL	dU
0.12	0.0026	0.0509	0.025	0.02	0.22

(3)　区間推定

このデータについて割合の差の点推定 $d = 0.12$ だけでなく，区間推定をしたい．

仮説検定と同様に正規近似を用いる．

そのためには，割合の差の標準誤差を求める必要がある．仮説検定のときには，帰無仮説が正しいという仮定の下での標準誤差が用いられた．具体的には，式(3.4)で $\pi_1 = \pi_2 = \pi$ として式(3.5)を用いた．

区間推定の場合には，仮説がないので，式(3.4)で，π_1，π_2 の代わりに p_1，p_2

を用いる．

$$\begin{aligned} \mathrm{V}[d] &= \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} \\ &= \frac{0.30(1-0.30)}{100} + \frac{0.18(1-0.18)}{300} = 0.0026 \\ se(d) &= \sqrt{0.0026} = 0.0509 \end{aligned}$$

割合の差 d に上に求めた差の標準誤差 $se(d)$ の 1.96 倍を加減すると，信頼率 95% の信頼区間が求められる．

$$0.12 - 1.96 \times 0.0509 = 0.02 < \pi_1 - \pi_2 < 0.22 = 0.12 + 1.96 \times 0.0509$$

Excel による計算結果を表示3.12 の下1/3 に示す．

(4) 例2

n の小さいデータの例として，新薬開発過程で得られたデータ（仮想データ）を取り上げる．

従来の処方では，15 例中有効は 4 例に過ぎなかったが，新しい処方では，10 例中有効が 7 例あった．すなわち，有効率は $4/15 = 27\%$ が $7/10 = 70\%$ に 43 ポイント向上した．これから新しい処方は有効率を向上するといえるであろうか．

表示3.12 の計算表に，この値を入力すると，表示3.13 が得られる．

この例では，新しい処方が有効率を低下することは考えられないので，片側の区間推定をすると，少なくとも有効率が 13 ポイント以上向上したということができる．

この例のように，n が小さいとき，正規近似では精度が低い．もっと正確な検定が必要となる場合は，§4.1 で説明する方法が用いられる．

■ **演習 15** 例2で，旧法の有効数の 4 を 6 とすると，結果はどうなるか？ ■

表示3.13: 2つの割合の差の検定と区間推定（例2）

	n	x	p
旧法	15	4	0.27
新法	10	7	0.70
合計	25	11	0.44

仮説検定

d	V[d]	se(d)	z	p値(片側)	p値(両側)
-0.43333	0.0411	0.2026	-2.138	0.016	0.032

区間推定

d	V[d]	se(d)	α(片側)	dL	dU
-0.43333	0.0340	0.1845	0.05	-0.74	-0.13

本日のまとめ

今日は2組の割合データがあるとき，割合の違いについて検定や推定をする方法を学んだ．ここで説明した正規近似による方法は，n がある程度大きいときには，この近似法で十分であり，広く用いられている．

適用範囲を拡張した 分割表による検定を §4.1 で説明する．§4.1 を読んでから，今日の内容をもう一度読み直すとより理解を深めることができるであろう．

4 その他の離散分布

4.1 多項分布と分割表

(1) 例

これまでに学んだ2項分布は，yes，no のいずれかの反応がある場合を扱うものであった．それに対して，3 つ以上の項目の中から一つを選択する場合がある．

「与党を支持する」,「野党を支持する」,「いずれも支持しない」のどれかを選んでもらうのも一つの例である．この場合，与党の支持率，野党の支持率を求める方法もあるが，無党派層の増減にも関心があるので，不十分である．

3つ以上の項目の割合を表わす分布が 多項分布 である．

(2) 多項分布

どの個体も 2 つの属性を持っているとき，それらの属性によって分類して作った表をクロス表と呼んだ (⇒第2 単元)．この表から，2 つの属性の間に関係があるかどうかを見るのであるが，その表が標本について作られている場合は，それが抽出された母集団で本当に関係を持っているかどうかを統計的に確かめる必要がある．そのために χ^2（カイ 2乗）検定 が用いられる．そのような検定をするとき，クロス表 を特に 分割表 と呼ぶ．

表示4.1 は所得階層と使用化粧品の銘柄との関係を調べるための調査結果である．表側には所得階層が表頭には化粧品の銘柄が取ってある．また，表の中の数字は人数である．

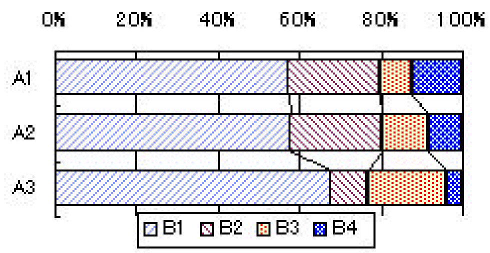
所得階層を i，銘柄を j，人数を f_{ij} で表わし，行数を a，列数を b とする．また，i 行目の計を $F_{i\cdot}$，j 列目の計を $F_{\cdot j}$，総計を $F_{\cdot\cdot}$ で表わすことにする．

表示4.2のようにグラフ化してみると，所得が高くなると B_3の化粧品 が減り，B_2, B_4 の化粧品 が増える傾向が見られる．このデータから，所得階層に

表示4.1: 分割表

			銘柄				
			B_1	B_2	B_3	B_4	行計
所得	上	A_1	$93(f_{11})$	$37(f_{12})$	$13(f_{13})$	$20(f_{14})$	$163(F_1.)$
	中	A_2	$241(f_{21})$	$96(f_{22})$	$48(f_{23})$	$26(f_{24})$	$421(F_2.)$
	下	A_3	$86(f_{31})$	$12(f_{32})$	$25(f_{33})$	$5(f_{34})$	$128(F_3.)$
	列計		$420(F._1)$	$145(F.2)$	$86(F._3)$	$61(F._4)$	$712(F..)$

表示4.2: 分割表のグラフ化



よって化粧品の銘柄選択が違うといえるだろうか.

(3)　期待度数

全体で，B_4の化粧品を使っている割合を求めると，$61/712 \times 100 = 8.57\%$ である．したがって．所得階層と化粧品の銘柄の間に関係がないならば，所得の高い階層でB_4を使う人は，$163 \times 0.0857 = 14.0$人になるはずである．これは，もし両者が無関係ならば，こうなるであろうと期待される度数であるから，これを 期待度数 と呼ぶ.

期待度数（$\widehat{f}_{ij}$ で表わす）は次の式で求められる.

$$\widehat{f}_{ij} = \frac{F_i.F._j}{F..} \tag{4.1}$$

$$\widehat{f}_{14} = \frac{163 \times 61}{712} = 14.0$$

となる.

同様にして，各組合わせの期待度数を計算すると，表示4.3の 期待度数の表が得られる．(列計と行計は，実測度数の計に一致する)

表示4.3: Excel による計算

		銘柄				
所得	度数	B1	B2	B3	B4	計
上	A1	93	37	13	20	163
中	A2	241	96	48	36	421
下	A3	**86**	12	25	**5**	128
	計	420	145	86	61	712

期待度数	B1	B2	B3	B4	計
A1	96.2	33.2	19.7	14.0	163
A2	248.3	85.7	50.9	36.1	421
A3	**75.5**	26.1	15.5	**11.0**	128
計	420	145	86	61	712

相対的外れ	B1	B2	B3	B4
A1	-0.32	0.66	-1.51	1.61
A2	-0.47	1.11	-0.40	-0.01
A3	**1.21**	-2.76	2.43	**-1.80**

a	3		
b	4	自由度	6
x^2	25.21	p値	0.0003

実際の度数(実測度数)が期待度数から外れているかどうかは「実測度数−期待度数」で測れるであろうか．A_3B_1 の差 $86-75.5=10.5$ を A_3B_4 の差 $5-11=-6$ と比べると前者の方が大きい．しかし，相対的に見ると，75.5 に対する 10.5 の差は 11 に対する 6 の差 よりも小さいと感じられる．この点を配慮して，実測度数と期待度数の隔たりの大きさを測るものさしとして

$$z_{ij}=\frac{f_{ij}-\widehat{f}_{ij}}{\sqrt{\widehat{f}_{ij}}} \tag{4.2}$$

$$z_{14}=\frac{20-14.0}{\sqrt{14.0}}=1.61$$

が用いられる．

具体的な値を表示4.3の 相対的外れ の表に示す．

z_{34} の絶対値は z_{31} よりも大きくなった．

最初に「所得が高くなると B_3の化粧品 が減り，B_2, B_4の化粧品 が増える傾向が見られる」と記したが，表示4.3 はこの傾向を示している．

(4)　χ^2 検定

所得と銘柄選択が本当は無関係（独立）であったとしても，z_{ij} が 0 になるわけではない．0 を中心としてばらつく．

z_{ij} を総合して，0 からの隔たりが誤差の範囲内かどうかを判断するために，2乗して加えた値が用いられる．$\pm$ の値を取る量を総合するために統計手法が用いる常套手段である．

このようにして導かれた値を χ^2 で表わす．

$$\chi^2 = \sum \frac{(\text{実測度数} - \text{期待度数})^2}{\text{期待度数}} = \sum_{i=1}^{a}\sum_{j=1}^{b} z_{ij}^2 \tag{4.3}$$

$$= 25.21 \tag{4.4}$$

この値が表示4.3 のχ ^2のセルに求められている．

ここで χ^2 という記号を使ったのは，第4単元で学ぶカイ2乗分布が関係するからである．ここではカイ2乗分布に深く立ち入らないで，上に得られた $\chi^2 = 25.21$ が誤差の範囲内かどうかを判断する手順だけを説明する．

χ^2 は $ab = 3 \times 4 = 12$ 個の z_{ij} を総合化したものであった．この計算の元になった $f_{ij} - \widehat{f}_{ij}$ は，横に加えても，縦に加えても 0 となる（$\widehat{f}_{ij}$ の行計，列計が $F_{i\cdot}$, $F_{\cdot j}$ に等しいことを思い出してほしい）．したがって，最後の行と最後の列が消えていても，残りの要素から埋めることができる．すなわち,「ab 個の要素の内，自由に決められる要素の個数は $(a-1)(b-1) = 2 \times 3 = 6$ である」．第2単元の§2.5(3) で学んだ 自由度 と同じ考えで，以下の計算で 自由度 を 6 とする．

式(4.3) で計算された χ^2 は,「分割表で行と列が独立である」という帰無仮説が正しければ自由度が $(a-1)(b-1)$ のカイ2乗分布に従い，自由度（ここでは6）からそれほど大きくなることは稀である．稀であるかどうかは，自由度6のカイ2乗分布が25.21 よりも大きい確率（p値）で判断される．この確率は 表

示4.3のχ＾2のセルに

`=CHIDIST(`χ^2 , 自由度)

として求められる.

この場合のp値は0.0003であって，この値は極めて小さいので，所得階層によって銘柄の選択が異なることは統計的に明白であるといえる.

■ 演習16

第2単元の表示1.13には業種によるパソコンの普及状況の違いの調査結果が示されている.

このデータを本日学んだ方法で解析し，業種によるパソコン普及状況の違いが統計的に有意かどうかを判断せよ.

ヒント テキストの計算表を使って，別の数値について解析したいときは，計算表をコピーして新しいデータを入力する．このとき，計算表の中でセル名に $ が入っていると，正しく計算されなくなる．これを避けるために，次の注意が必要である.

新規のシート上で計算したいときは，計算表を新シート（新規のブックでも可）にコピーする．その際，元の計算表の位置（行番号と列名）が同じところになるように注意する．シートの左上などに移動したいときは，不要の行や列を削除する.

既存のシートの希望する位置にコピーしたいときは，上のように新しいシート上にコピーし，位置を修正してから，既存のシートにコピーする. ■

本日のまとめ

今日学んだ分割表の χ^2検定は，応用範囲の広い手法である．第2単元の §1 でいろいろのクロス表の作り方を学んだ．そこでは，単に割合が異なることをグラフ化するだけであったが，割合の違いが統計的に意味があるかどうかを判断するためには，分割表が用いられる.

4.2　2 × 2 分割表とその展開

上に説明した分割表を 2 行2列 にすれば，§3.4 で説明した，2 つの割合の違いの検定に適用できることは容易に想像される．

このような分割表を **2 × 2 分割表** と呼ぶ．

(1)　2 × 2 分割表

§3.4 で取り上げた2つの例で説明する．

実測度数から，表示4.4 の上に示す2 × 2 分割表を作成する．

表示4.4: 2 × 2 分割表

例1

実測度数	喫煙	非喫煙	計
患者	30	70	100
非患者	54	246	300
計	84	316	400

期待度数	喫煙	非喫煙	計
患者	21.0	79.0	100
非患者	63.0	237.0	300
計	84	316	400

相対外れ	喫煙	非喫煙	計
患者	1.96	-1.01	4.882
非患者	-1.13	0.58	1.627
計	5.143	1.367	6.510
p値			0.011

例2

実測度数	有効	無効	計
旧法	4	11	15
新法	7	3	10
計	11	14	25

期待度数	有効	無効	計
旧法	6.6	8.4	15
新法	4.4	5.6	10
計	11	14	25

相対外れ	有効	無効	計
旧法	-1.01	0.90	1.829
新法	1.24	-1.10	2.744
計	2.561	2.012	4.573
p値			0.032

最初の分割表から，表示4.3 と同様の手順で，χ^2 とそのp値を求めるまでの過程を一つの計算表にまとめてある．

帰無仮説が正しいとき，χ^2 は，近似的に自由度が $(a-1)(b-1) = 1 \times 1 = 1$ のカイ 2 乗分布に従う．

こうして得られた2つの例のp値を，表示3.12（p.74），表示3.13（p.76）のp値（両側）と比較すると，

	§3.4		§4.2	
	z	p値	χ^2	p値
例1	2.551	0.011	6.510	0.011
例2	-2.138	0.032	4.573	0.032

となる．

p値は2つの方法で一致している．

z と χ^2 を比較すると，$\chi^2 = z^2$ の関係のある[1]．

これから，分割表の χ^2検定 は 両側検定 になっていることが分かるであろう．

2×2分割表 の場合は，χ^2 を次の式で直接求めることができる．

$$
\begin{aligned}
\chi^2 &= \frac{(f_{11}f_{22} - f_{12}f_{21})^2 F_{\cdot\cdot}}{F_{1\cdot}F_{2\cdot}F_{\cdot 1}F_{\cdot 2}} \qquad (4.5)\\
&= \frac{(30 \times 246 - 70 \times 54)^2 \times 400}{100 \times 300 \times 84 \times 316} = 6.510 \quad \text{例1}\\
&= \frac{(4 \times 3 - 11 \times 7)^2 \times 25}{15 \times 10 \times 11 \times 14} = 4.573 \quad \text{例2}
\end{aligned}
$$

この式を用いて解析するExcelシートを表示4.5に示す．

表示4.5: 2×2分割表 (2)

	A	B	C	D	E	F	G	H	I
24		喫煙	非喫煙	計			有効	無効	計
25	患者	30	70	100		旧法	4	11	15
26	非患者	54	246	300		新法	7	3	10
27	計	84	316	400		計	11	14	25
28									
29	χ2	6.510		5.807		χ2	4.573		2.983
30	p値	0.0107		0.0160		p値	0.0325		0.0841
31									
32	B29:	=(B25*C26-C25*B26)^2*D27/(D25*D26*B27*C27)							
33	B30:	=CHIDIST(B29,1)							
34	D29:	=(ABS(B25*C26-C25*B26)-D27/2)^2*D27/(D25*D26*B27*C27)							

表示4.5の χ^2（B29, G29） は式(4.5)を用いて計算されている（D, I列の数値については次項で説明する）．

[1] 代数的に等しい値になることは数式を操作することで示すことができる．ここには省略する．

上に示した方法のどれを用いても同じ結果が得られる．計算の簡便さからは式(4.5)が便利である．

(2)　連続修正

2項分布の正規近似で，x が整数のみをとるために 連続修正 を加えることで正確な2項分布に接近できることを説明した．

2*2分割法で連続修正を追加する方法がイェーツによって提案された．

連続修正とは，式(4.5) の分子の2乗を求める値 $f_{11}f_{22} - f_{12}f_{21}$ の絶対値から $F../2$ を引くものである．

上の2つの例について計算すると，

$$\begin{aligned}\chi^2 &= \frac{(|f_{11}f_{22} - f_{12}f_{21}| - F../2)^2 F..}{F_1.F_2.F._1F._2} \qquad (4.6)\\ &= \frac{(|30 \times 246 - 70 \times 54| - 400/2)^2 \times 400}{100 \times 300 \times 84 \times 316} = 5.807\\ &= \frac{(|4 \times 3 - 11 \times 7| - 25/2)^2 \times 25}{15 \times 10 \times 11 \times 14} = 2.983\end{aligned}$$

となる．この値が表示4.5 の D29, I29 に求められている．

連続修正を加えない結果と比較すると，f_{ij} が大きい例1ではχ^2 が 6.510 が 5.807 になり，p 値が 0.011 が 0.016 になる程度の違いで実質的な差はない．それに対して，例2では χ^2 が 4.573 が 2.983 になり，p 値が 0.032 が 0.084 になる．すなわち，5%有意が有意ではなくなる．

以上の数値例から，f_{ij} が小さいときは連続修正を加えるべきだと思われる．しかし，§2.5(2) で,「2 項分布の正規近似による検定で連続修正を加えると実質的な第1種の過誤の確率が小さくなる」という事実と同様な理由から，多くの場合連続修正を加えない方法で十分実用に使えると考えられる．

(3)　正確な方法

前項の最後に説明した連続修正によって正規近似の精度を改善することができるが，フィシャーは正確な方法を提案した．

この方法は フィッシャー の確率計算法 と呼ばれる．

周辺の度数（$F_{1\cdot}$, $F_{2\cdot}$, $F_{\cdot 1}$, $F_{\cdot 2}$）を固定したとき，f_{11}，f_{12}，f_{21}，f_{22} となる確率は，1行目の合計が $F_{1\cdot}$ で，左のセルの値が f_{11} である組合わせの個数 ${}_{F_{1\cdot}}C_{f_{11}}$ と，2行目についての同様の組合わせ個数 ${}_{F_{2\cdot}}C_{f_{21}}$ の積を，合計の行について同様の組合わせ個数 ${}_{F_{\cdot\cdot}}C_{F_{\cdot 1}}$ で割って求められる[2]．式で表わすと，

$$
\begin{aligned}
\text{確率} &= \frac{{}_{F_{1\cdot}}C_{f_{11}} \cdot {}_{F_{2\cdot}}C_{f_{21}}}{{}_{F_{\cdot\cdot}}C_{F_{\cdot 1}}} = \frac{F_{1\cdot}!}{f_{11}!\, f_{12}!} \times \frac{F_{2\cdot}!}{f_{21}!\, f_{22}!} \Big/ \frac{F_{\cdot\cdot}!}{F_{\cdot 1}!\, F_{\cdot 2}!} \\
&= \frac{F_{1\cdot}!\, F_{2\cdot}!\, F_{\cdot 1}!\, F_{\cdot 2}!}{F_{\cdot\cdot}!\, f_{11}!\, f_{12}!\, f_{21}!\, f_{22}!} \qquad (4.7) \\
&= \frac{F_{1\cdot}!\, F_{2\cdot}!\, F_{\cdot 1}!\, F_{\cdot 2}!}{F_{\cdot\cdot}!} \times \frac{1}{f_{11}!\, f_{12}!\, f_{21}!\, f_{22}!} \qquad (4.8)
\end{aligned}
$$

となる．

例2の分割表について計算すると，

$$
\text{確率} = \frac{15!\, 10!\, 11!\, 14!}{25!\, 4!\, 11!\, 7!\, 3!} = 0.0367
$$

となる．

2項分布の検定で $x = 2$ のとき，$x = 2$ の確率だけでなく，$x = 1$, 0 の確率も加えてp値を求めたのと同様に，この場合も，得られた2元表の確率だけでなく，もっと新法が良い結果の得られた確率も求めて，合計する必要がある．

そのための計算表が表示4.6である．

式(4.8)で確率を計算するために，共通の最初の比率 $\frac{F_{1\cdot}!\, F_{2\cdot}!\, F_{\cdot 1}!\, F_{\cdot 2}!}{F_{\cdot\cdot}!}$ を`O11`に計算しておく[3]．

`K14:N14`に分割表の度数をコピーする．`O14`に上の確率 0.03675 が求められている．

f_{11} の実測度数がその期待度数よりも大きいか・小さいかによって，f_{11} の度数をどちらに変化させるかが決まる．$f_{11}\, f_{22} < f_{21}\, f_{12}$ であれば 小さい方に，逆であれば大きい方に変化させる．この例の場合は，$f_{11}\, f_{22} = 12$, $f_{21}\, f_{12} = 77$ であるから，小さい方に変化させることになる．

---

[2] この式がどのようにして導かれたかの説明は省略する．

[3] 総数 $F_{\cdot\cdot}$ が170を超えると，Excel の `FACT` 関数がオーバーフローを起こすため $F_{\cdot\cdot}!$ が計算ができない．$F_{\cdot\cdot}$ が大きい場合には近似法で十分である．

表示4.6: フィッシャー法の計算過程

	K	L	M	N	O	P
5		有効	無効	計		
6	旧法	4	11	15		
7	新法	7	3	10		
8	計	11	14	25		
9						
10				変化させる方向	-1	
11				F1.*F2.*F.1*F.2/F..	1.06E+12	
12						
13	f11	f12	f21	f22	Prob	累積
14	4	11	7	3	0.03675	0.03675
15	3	12	8	2	0.00459	0.04134
16	2	13	9	1	0.00024	0.04158
17	1	14	10	0	0.00000	0.04158
18						
19	O10: =IF(L6*M7<M6*L7,-1,1)					
20	O11: =FACT(N6)*FACT(N7)*FACT(L8)*FACT(M8)/FACT(N8)					
21	O14: =O$11/(FACT(K14)*FACT(L14)*FACT(M14)*FACT(N14))					
22	K15: =K14+O10	L15: =L14-O10				

それが f_{11} を変化させる方向 O10 に求められており，-1 となっている．

f_{11}，f_{22} を減少し，f_{12}，f_{21} を増加する．どこかのセルが 0 になるまで，下にコピーする．

各行について確率と，その累積を計算する．

確率を合計した 0.04158 が，片側 p値である．したがって，2つの方法を比較する場合は，フィッシャーの方法のp値を2倍して比較しなければならない．

もし，新法が旧法よりも良いことだけを確認したいのであれば，片側検定である．片側p値が 0.05 以下であるから，新法は旧法よりも有効率が向上したといえる．

■　演習 17　　ある商品の名前を知っている人の割合（認知率）がテレビのコマーシャルによって増えたかどうかを確認するために調査を行った．

放送前に，ある地点の通行人100人に聞いたところ，29人が知っていた（$n_1 = 100$，$x_1 = 29$，$p_1 = 0.29$）．

新しいコマーシャルを放送後同様の調査を実施し，$n_2 = 150$，$x_2 = 63$，$p_2 =$

0.42 が得られた．

このデータについて，表示3.12（p.74）の方法と，表示4.5（p.83）の方法で解析し，両者の結果を比較せよ． ■

本日のまとめ

今日は，昨日勉強した分割表を，2 つの割合の比較に利用する方法を学んだ．結果として，§3.4 で学んだ正規近似と同じ結果が得られ，実用的には無意味であったかもしれない．しかし，別の考えから同じ解析方法が導かれることを知ることは，統計について勉強する上で役に立つであろう．

最後に連続修正とフィッシャーの正確な方法を学んだ．この方法は，サンプルが比較的小さく，フォーマルな解析を要求される場で役に立つことがあるであろう．

4.3 ポアッソン分布

(1) 稀に起こる現象

生徒数 $n = 50$ 人の学級で，日々の欠席者数 x の統計を取っているとする．流感におそわれたというような時を別にすると，日々の欠席率 π は一定と仮定できる．すると，確率 π で起こる事象が，大きさ $n = 50$ の標本で起こる回数 x の分布を考えることになり，それは，§2 の 2 項分布となる．ところで，この場合欠席するというのは稀な現象であって，欠席する生徒の確率 π は非常に小さいのが普通である．したがって，日々の統計で $x = 0, 1, 2$ というような値はしばしば起こるが，$x \geq 10$ つまり 10 人以上の生徒が同じ日に休むというような事象は，めったに起こらない．このようなときに x の取り得る値として，0，1，2，⋯，n の全部を考えるのは大変無駄である．

第 2 単元§3.1 で取り上げた，ある都市の交通事故による毎日の死亡者数の統

計は，多い日で 9 人，少ない日は 0 人になっている．この現象でも，人は交通事故によって死亡するか否かの 2 つの場合があり，その死亡率 π が考えられるが，2 項分布とするためには，標本の大きさ n を指定しなければならない．この n としてその市の住民の総数を考えるのは適当でない．市民の中でもその日一歩も外出しなかった人は，交通事故にあう可能性はないし，よそから来てその市内で交通事故にあう人もある．だから，n は毎日変わると考えられる．この死亡率 π は，何百万人のうちの 1 人とか 2 人という割合であるから，交通事故で死ぬことは非常に稀な現象であるといわねばならない．このような場合には，n を無視して，交通事故で死亡した人の数 x だけに注目すれば良い．

(2)　2 項分布の極限としてのポアッソン分布

事象 E は出現率 π の 2 項分布に従って生起するとき，大きさ n の標本における E の出現数 x の期待値，すなわち，2 項分布の平均は式 (2.4) (p.34) より $\mu = n\pi$ であった．この μ を固定したままで，n を十分大きくし，それに応じて π を十分小さくしたときの極限分布を数学的に求めると（その導出はここではやらない），次の確率で表わされる **ポアッソン分布** になる．

$$p_x = e^{-\mu}\frac{\mu^x}{x!} \qquad x = 0, 1, 2, \ldots \tag{4.9}$$

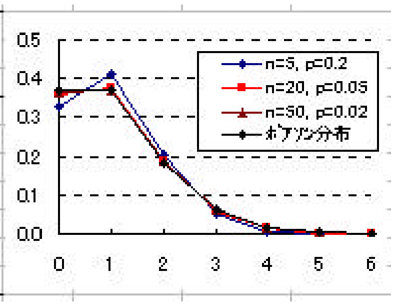
n を無限に大きくしたのだから，x も $0,\ 1,\ 2,\ \cdots$ と無限に大きくなり得るが，x がすこし大きくなると，p_x はほとんど 0 になる．ここで，e は自然対数の底で $2.71828\cdots$ である．この式から，ポアッソン分布の確率 p_x は，平均 μ のみによって定まることが分かる．

大きさ n，出現率 π の 2 項分布の確率が，平均が $\mu = n\pi$ に等しいポアッソン分布の確率にどのくらい近いかを表示 4.7 に示す．

ここでは，$\mu = 1$ のポアッソン分布を，$n = 50$，$\pi = 0.02$; $n = 20$，$\pi = 0.05$; $n = 5$，$\pi = 0.20$ の 3 つの 2 項分布と比較した．$n = 50$ の 2 項分布の確率とは，小数点以下第 2 位まで一致する．$n = 20$ の 2 項分布の確率もポアッソン分布の確率にかなり近い．

表示4.7: 平均 μ の等しい2項分布とポアソン分布の確率の比較

x	2項分布			ポアソン分布
	5	20	50	
	0.2	0.05	0.02	1.0
0	0.328	0.358	0.364	0.368
1	0.410	0.377	0.372	0.368
2	0.205	0.189	0.186	0.184
3	0.051	0.060	0.061	0.061
4	0.006	0.013	0.015	0.015
5	0.000	0.002	0.003	0.003
6		0.000	0.000	0.001



しかし，$n = 5$, $\pi = 0.20$ の2項分布は，表示4.7で分かるように，ポアソン分布の確率とはかなり食い違う．

(3) μ によるポアソン分布の変化

Excel にはポアソン分布の確率と累積確率を計算する関数が準備されている．例えば，表示4.7で，$\mu = 1$, $x = 2$ の確率(0.184) は

```
=POISSON(x, μ,FALSE) =POISSON(2,1.0,FALSE)
```

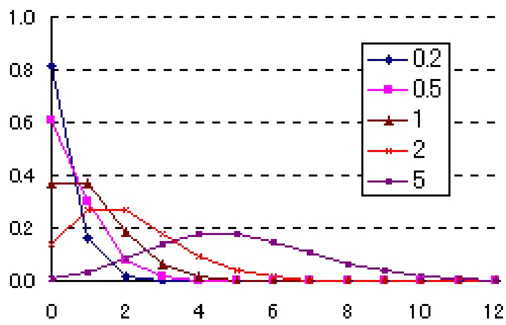
で求められ，最後の FALSE を TRUE とすると累積確率(0.368+0.368+0.184=0.920) が求められる[4]．

この関数を使って，平均 μ の異なる種々のポアソン分布の確率 p_x を計算し，グラフ化した結果を表示4.8 に示す．

μ の小さいポアソン分布では，その確率 p_x は $x = 0$ のとき最大で，x が大きくなるとともに減少する．μ が大きくなるにつれて分布に山ができ，その山はだんだん右へ移動するとともに低くなる．μ が 5.0 ぐらいになると，分布はほぼ左右対称の形となる．2項分布の場合と同様，$\mu > 5$ のポアソン分布は，正規分布で近似することができる．

[4] x, μ が大きい値のときは，計算不能となる．そのときは，後に述べる正規近似を使う．

表示4.8: 平均 μ が異なるポアッソン分布



(4)　期待値と分散

ポアッソン分布は2項分布で，$n\pi = \mu$ を固定して，$\pi \Rightarrow 0$ とした極限であった．

2項分布の期待値と分散の式で，$n\pi$ を μ，π を 0 に置き換えると，ポアッソン分布の期待値と分散の式が導かれる．

$$\mathrm{E}[x] = n\pi = \mu \tag{4.10}$$

$$\mathrm{V}[x] = n\pi(1-\pi) = \mu \tag{4.11}$$

これから，ポアッソン分布では，期待値と分散が等しいことが分かる．

■　演習18　新聞輪転印刷機で紙が切れて停止する事故は，1日に平均2回起こるとする．

(i)　事故が5回起こる日はどのくらいあるか．

(ii)　6回以上の事故が起こる日はどのくらいあるか．

(iii)　事故が1回も起こらない日の割合はいくらか．　■

(5) 点推定

織物の生産工程で，従来は織物1本当たりの平均欠点個数は3個であった．今回，工程に改善を加えた．改善後に生産されたたくさんの織物の集りから10本の織物を無作為に抜き取り，欠点(織りキズ)の個数を調べたところ，次のデータを得た．

1，0，2，2，4，3，1，3，5，1

これから，もとの織物全体での織物1本当たりの平均欠点個数を推定したい．

標本についての平均個数で母集団での平均個数を推定することができる．すなわち，

$$\begin{aligned}\overline{x} &= (1+0+2+2+4+3+1+3+5+1)/10\\ &= 2.2\end{aligned}$$

が平均欠点数 μ の点推定値となる．

(6) 仮説検定

このデータから，工程改善の効果があったといえるであろうか．

もし，改善の効果がなければ，10本の織物には $3 \times 10 = 30$ 個前後の欠点が含まれることが予想される．

実測された22個と30個との差が，統計的に意味のある差であるかどうかで判断すれば良い．

これは，帰無仮説 $\mathrm{H}_0 : \mu = 30$ の下で，$T = 22$ が起こりにくい（起こる確率が5%以下である）かどうかを調べることにより，分かる．ここでは，改善効果があることを確かめるのが目的であるから，片側検定となる．

$\mu = 30$ のポアッソン分布で，22以下の確率は，

```
=POISSON(22,30,TRUE)
```

で計算され，0.081 であることが分かる．

2項分布の正規近似のときと同様に 0.5 の補正を行い，正規近似すると，

$$z = \frac{|22-30|-0.5}{\sqrt{30}} = 1.369$$

すなわち，22は30から1.369σ離れている．標準正規分布で，1.369 以上の確率は 0.085 となり，上の正確な値 0.081 とほぼ等しい．この値は 0.05 よりも大きいから，これだけのデータだけでは，工程改善の効果があったことは確認できない．

もし，15本の織物を調べ平均欠点数が同じく2.2 であったとすると，実測度数 $2.2 \times 15 = 33$ と期待度数 $3 \times 15 = 45$ との差を検定すれば良い．

$\mu = 45$ のポアッソン分布で33以下の確率は0.038 である．これならば，改善効果のあったことが統計的に確かめられたことになる．

以上の計算を表示4.9 の上半分に示す．

表示4.9: ポアッソン分布の仮説検定と区間推定

	μ	n	nμ	T	下側確率	上側確率
仮説検定	3	10	30	22	0.081	0.946
	3	15	45	33	0.038	0.974
区間推定	3.33	10	33.3	22	**0.025**	0.984
	1.38	10	13.8	22	0.986	**0.025**
	3.14	10	31.4	22	**0.050**	0.968

この試算が示すように，仮説検定で「有意差がない」というのは,「差がない」のではなく,「差のあることの確証が得られない」という意味である．データ数を増やすことにより，確証が得られる可能性が残っている．とくに，p値がα（例えば 0.05）よりもわずかに大きいときは，その可能性が高い．

(7)　区間推定

2項分布の場合と同様の手順で μ の信頼区間を求めることができる．

すなわち，表示4.9の下半分に示すように，下側確率または上側確率がちょうど0.025 になる μ の値を探索的に，または，ゴールシークを使って，求める．

この例では，反物1本当たりの平均欠点数は 1.38 から 3.33 の間であるという，両側信頼率が95%の信頼区間が求められる．

もし，平均欠点数は**以下であるという区間推定をしたい場合には，下側確率

が 0.05 となる μ を求めれば良い．表示4.9の一番下の行に示すように，$\mu \leq 3.14$ という片側の信頼区間が得られる．

ポアッソン分布の場合，次の式を使って信頼区間 μ_L，μ_U を直接計算することができる．

$$\mu_L = \chi^2(1 - \alpha/2;\ 2x)/2 \tag{4.12}$$

$$\mu_U = \chi^2(\alpha/2;\ 2(x+1))/2 \tag{4.13}$$

ここに，$\chi^2(\alpha/2,\ \ldots)$ は前節で出てきたカイ 2 乗分布の上側パーセント点である．Excel では，

```
=CHIINV(1-α/2, 2*x)/2
=CHIINV(α/2, 2*(x+1))/2
```

と書ける．

これも，2 項分布の場合と同じように，上記の計算式を VBA に登録しておくことにより容易に利用できる．テキストに添付されたファイルには

```
=P_PoissonL(X, P2), =P_PoissonU(X, P2)
```

として記録されている．使い方は2項分布の場合と同じである．

(8)　ポアッソン分布かどうかの検定

前節でも引用した第2単元§3.1の表示3.1のデータを再び用いる．その $m = 305$ 日にわたる度数分布は，やはり第2単元§3.1の表示3.2に示されているが，記号を一部変えて，次の表示4.10に再録する．これから，交通事故による1日当たり死亡者数の平均は，

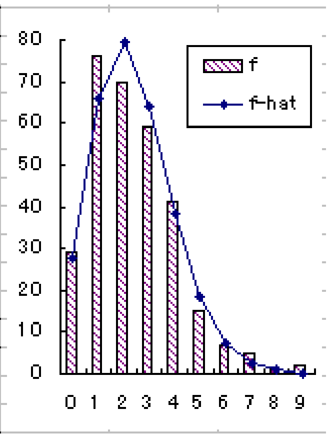
$$\overline{x} = 735/305 = 2.410 \text{ 人}$$

であることが分かる．

いま，この $\overline{x} = 2.410$ をポアッソン分布の平均 μ であると考えて，各 x に対するポアッソン分布の確率を POISSON 関数を使って計算し，それに$m = 305$ を掛けて f の期待度数を求める．その結果を表示4.10左の表の $\widehat{f}$ の欄に書きこむ．

表示 4.10: 交通事故死亡者数の度数分布とポアッソン分布の当てはめ

	f	f-hat
0	29	27.40
1	76	66.03
2	70	79.55
3	59	63.90
4	41	38.50
5	15	18.56
6	7	7.45
7	5	2.57
8	1	0.77
9	2	0.21
合計	305	
平均	2.410	
平方和	869.770	
平均平方	2.861	



f（柱）と $\widehat{f}$（折れ線）の関係が右のグラフに示されている．

データにはポアッソン分布が良く合っているように見える．しかし，注意深く眺めると，f_1 が大きく，また x の大きい右の裾の柱が折れ線の上に飛び出している．

これからポアッソン分布よりもバラツキが大きく，長い裾を引いているのではないかと思われる．

そこで，もとの度数から直接平方和と平均平方を計算すると，

$$S = 869.770$$
$$V = 2.861$$

が得られる．

ポアッソン分布は期待値と分散が等しい．それに対してこのデータは平均値が 2.410 に比べて平均平方が 2.861 と幾分大きい．この違いは統計的に意味があるかどうかを検定する方法は，第 4 単元で詳しく説明される．ここでは，その結果だけを示す．第 4 単元の学習が終わったらもう一度読み直してほしい．

平方和が真の分散（ここでは，ポアッソン分布を仮定し，平均値を用いる）の何倍あるかを求め，χ^2 で表わす．

$$\chi^2 = \frac{869.770}{2.410} = 360.925$$

もし，平均値と平均平方が等しければ，この比は自由度 $305 - 1 = 304$ に近い値を取るはずである．

ここに得られた 360.925 は 304 に比べて，統計的に意味のある差があるかどうかは，自由度が 304 のカイ 2 乗分布で，360.925 以上の確率（p値）で判断される．この確率は 0.0137 で，これは 0.05 よりも小さい．したがって，交通事故による死亡者の分布は，ポアッソン分布で予想されるよりもバラツキが大きいといえる．

死亡者数の分布がポアッソン分布よりも大きなバラツキを持つ理由は2つ考えられる．

- 事故が雨の日や月末に多く，日曜日に少ない．
- 事故件数はポアッソン分布に従うとしても，1回の事故で数人が死ぬ．

ポアッソン分布は，個々の事象が独立に発生するという前提で導かれるものであることを知っておいてほしい．

製造現場の欠点数のデータで同様の傾向が観察された場合には，データの層別などにより，製品の欠点数を減らすためのヒントが得られる可能性が高い状態といえる．

本日のまとめ

今日は，ポアッソン分布について学び，稀に起こる事象の回数の分布がポアッソン分布であり，2項分布のある極限として導かれることを知った．また，ポアッソン分布の確率や累積確率は，2項分布と同様に，Excel関数で簡単に計算できることも学んだ．

ついで，ポアッソン分布に従うことが分かっている場合の，検定や区間推定と，ポアッソン分布に従うかどうかを調べる方法を学んだ．

本文にも書かれているように，ポアッソン分布は，個々の事象が独立に起こることが前提である．流感などが流行して，複数の生徒が集団で欠席するとい

うような場面では，ポアッソン分布は当てはまらない．

4.4 幾何分布

(1) 例

2人でじゃんけんをすると，勝負の つくとき と つかないとき がある．2人とも，石，紙，はさみを出したときは勝負がつかないが，2人が違うものを出せば勝負がつく．各人は，石，紙，はさみを同じ確率で，ランダムに出すとする．起こり得る9通りの場合は同等で，いずれも $1/9$ の確率を持つ．この9通りのうち6通りは勝負がつき，3通りは勝負が決まらない．したがって，勝負のつく確率は $\pi = 6/9 = 2/3$ である．

このとき，2項分布では $n = 5$回じゃんけんをしたら，そのうちx 回勝負のつく確率を求めることができる．

ここで，見方を変えて，n を固定しないで，何回目に勝負がつくかを知る方法を考える．

5回じゃんけんをしてもまだ勝負がつかない確率を知りたい．または，勝負がつかない確率が5%以下になるのは，何回かを知りたい．

(2) 2項分布との違い

2項分布では，ある事象Eの起こる確率πが与えられていて，n回試行したときに，そのうちのx回にEの起こる確率p_xを定義した．これとは逆に，ある事象Eが起こってほしいと思っているが，何回目にはじめて起こるかを知りたい場合がある．このときには，nは固定されておらず，x回目に初めて事象Eが起こるとして，そのxを確率変数とする．x回目に<u>はじめて</u>Eが起こるということは，第1回目から第$(x-1)$回目まではEが起こらないことを意味するから，確率p_xは，

$$p_x = \pi(1-\pi)^{x-1} \qquad x = 1, 2, \dots \tag{4.14}$$

と書け，次の関係がある．

$$p_{x+1} = \pi(1-\pi)^x = \pi(1-\pi)^{x-1}(1-\pi) = p_x(1-\pi) \tag{4.15}$$

第0回目にEが起こるということはありえないから，xは1から始まり，際限なく大きくなりうる．よって，xのとりうる値は，$1, 2, \dots$としたのである．このp_xはちょうど，初項π，公比$1-\pi$の幾何数列（等比数列）の形をしているから，この分布を幾何分布と呼ぶ．

(3) 幾何分布の確率と累積確率

そこで，$\pi = 2/3$の幾何分布の確率と累積確率を計算したのが表示4.11の「じゃんけん」の欄である．

表示4.11: 幾何分布の確率と累積確率

	A	B	C	D	E	F	G	H	I	J
1		じゃんけん					空車			
2		π	1-π			π	1-π			
3		0.667	0.333			0.2	0.8			
4	x	確率	累積確率		x	確率	累積確率	x	確率	累積確率
5	1	0.667	0.667		1	0.200	0.200	8	0.042	0.832
6	2	0.222	0.889		2	0.160	0.360	9	0.034	0.866
7	3	0.074	0.963		3	0.128	0.488	10	0.027	0.893
8	4	0.025	0.988		4	0.102	0.590	11	0.021	0.914
9	5	0.008	0.996		5	0.082	0.672	12	0.017	0.931
10	6	0.003	0.999		6	0.066	0.738	13	0.014	0.945
11	7	0.001	1.000		7	0.052	0.790	14	0.011	0.956

表の上にはπと$1-\pi$が入力されている．

確率p_xは，式(4.14)を使っても計算できるが，上から逐次に計算する方が容易である．$x=1$の確率p_1にはπを入れる（B5: =B3）．$x>1$の確率は，上の確率に$1-\pi$を掛けて求める（B6: =B5*C$3）．B6を下にコピーする．

累積確率P_xは，確率を上から足し合わせたものであるが，次の式を使えば

計算することができる[5].

$$
\begin{aligned}
P_x &= p_1 + \ldots + p_{x-1} + p_x \\
(1-\pi)P_x &= (1-\pi)p_1 + \ldots + (1-\pi)p_{x-1} + (1-\pi)p_x \\
&= p_2 + \ldots + p_x + p_{x+1} \\
P_x - (1-\pi)P_x &= p_1 - p_{x+1} = \pi(1-(1-\pi)^x) \\
\pi P_x &= \pi(1-(1-\pi)^x) \\
P_x &= 1-(1-\pi)^x
\end{aligned}
$$

C5 のセルに `1-C$3^A5` を入力し，下にコピーする．

累積確率が 0.95 をはじめて超える x は 3 である．

つまり，じゃんけんでは,「3回までで勝負がつく」と95%以上の確からしさでいえるのである．

別の例を考える．ある街角では，空車は5台に1台の割合で来るとする．すなわち，空車である確率は $\pi = 1/5$ である．

じゃんけんと同様の計算を表示4.11の「空車」の欄に示す.

「5台待てば空車が拾える」確率は，表示4.11の $x = 5$ に対する累積確率から 0.672 であることが分かる．

すなわち，5台待てば約2/3の確率で空車が拾えるのである．

それでは，10台待っても空車が拾えないことはどのくらいの割合で起こるであろうか？ $x = 10$ までの累積確率が 0.893 であるから，10回に1回の割で10台待っても空車の拾えないことも起こる．

(4)　幾何分布の期待値と分散

幾何分布の期待値と分散は次の式で計算される（証明は省略）．

$$\text{期待値：}\mathrm{E}[x] = \frac{1}{\pi} \tag{4.16}$$

[5] 1行目の式に $1-\pi$ を掛けて2行目を求める．式(4.15)の関係を使うと，3行目の式が得られる．1行目から3行目を引くと4行目となり，整理すると最後の式が導かれる．

分散：$\mathrm{V}[x] = \dfrac{1-\pi}{\pi^2}$

これから，先の空車の例では$\pi = 1/5$より，

$$\mathrm{E}[x] = 5, \qquad \mathrm{V}[x] = (4/5)/(1/5)^2 = 20$$

を得る．よって，標準偏差$\sigma = \sqrt{20} = 4.47$となる．これは平均して，5台目には空車を拾うことができるが，標準偏差が4.47と大きいので，その標準偏差の2倍，3倍を加えると，10台，15台待っても空車が拾えないことがあるのを示している．表示4.11の累積確率を見てもそのことが分かる．

2項分布とポアッソン分布は正規分布に近づくが，幾何分布は正規分布で近似することができない．したがって，上に求めた期待値や標準偏差を使って判断するときには注意が必要である．

本日のまとめ

今日学んだ幾何分布の説明で取り上げた2つの例は，受講生の身の回りの問題として興味が持てたであろう．統計には多くの手法があり，本講座ですべてを紹介することはできないが，将来機会があれば，勉強してほしい．

4.5 補遺

(1) 離散量に関するまとめ

第3単元では離散分布の代表である2項分布について，たくさんの方法を説明した．

そこで取り上げられた手法を表にまとめた．

表示4.12: 2項分布に関する解析手法のまとめ

解析目的	解析手法	参照
割合 π の検定	正確法による検定	§2.4(1),(2)
	正規近似による検定	§2.4(3),(4)
割合 π の区間推定	正規近似（スコア法）による推定	§3.2(1)
	正確法による推定	§3.2(2)
サンプルサイズ		§3.3(3)
2つの割合の差の検定	正規近似による検定	§3.4(2)
	2*2分割表	§4.2(1)
	Fisher の確率計算法	§4.2(3)
2つの割合の差の推定	正規近似による推定	§3.4(3)

(2)　ギリシャ文字の使い方

サンプルから計算された割合には p を用いる．それに対して，母集団の割合には p に対応するギリシャ文字 π を用いる．

この例のように，サンプルから計算した値は アルファベット，母集団の値は ギリシャ文字 を用いるのが一般的である．

サンプルの標準偏差は s，母集団の標準偏差は σ で表わされる．

サンプルの平均値は $\overline{x}$ で表わしたが，mean の頭文字の m を使うこともある．これは，母集団の平均 μ に対応している．

第1種，第2種の誤りの確率には α，β が用いられている．

5　演習解答

5.1　第 1 章　仮説検定の基礎

演習 1 (p.11)

当たる確率が $1/3$ であるから，外れる確率は $1-1/3=2/3$ となる．6 枚が全部外れる確率は $(2/3)^6=0.088$ となる．この値はそれほど小さくない．

苦情を申し立てるためには根拠不十分であり，運が悪かったと諦めることになる．

枚数が $7, 8$ 枚のとき，全部外れる確率は $(2/3)^7=0.059$, $(2/3)^8=0.039$ となる．

8 枚が全部外れのときは，くじの公正さに疑問が残る．7 枚が全部外れのときは，判断にまようことになる．

演習 2(p.14)

A: 自動車を運転している．B: 免許証を持っている．

$A \Rightarrow B$ はほぼ正しい．

この関係はベン図を 1 次元にした次の図で表わされる．

```
運転を    |   している   |          していない          |
          |=============|-------------------------|
          |*|           |Paper Drv.|
          |*|=====================|--------------|
免許証を  |          持っている          |  持っていない  |
```

左の * の部分が無免許運転で，ほぼに相当する．

$A \Leftarrow B$ の成立しない部分が中央の Paper Drv.（ペーパードライバー）に相当する．

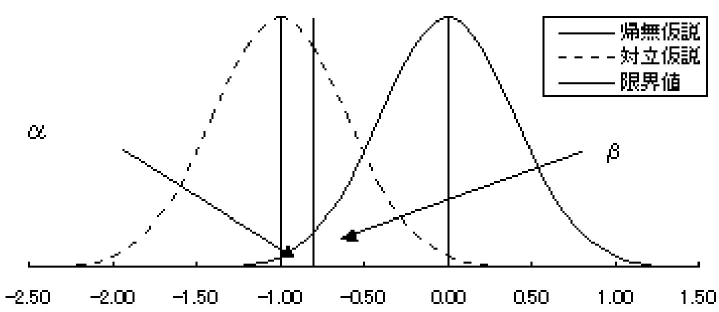
演習 3(p.22)

品質は「大きい方が良い」として考えを進める．

新原料または旧原料を使って生産した製品の品質の真の平均（母平均）を μ_{new}，μ_{old}，観測値の平均を $\overline{x}_{new}$，$\overline{x}_{old}$ とする．

$\delta = \mu_{new} - \mu_{old} = 0$ のとき，$d = \overline{x}_{new} - \overline{x}_{old}$ は 0 を中心とし，表示5.1の右の曲線のように分布する．

表示5.1: αとβの関係



対策(1) は「旧原料が有意に良いときに，旧原料を用いる」であるから，棄却域はこの分布の左側の確率が α となる領域である．これは

　　旧原料が新原料に比べ良くないのに，旧原料を使う

誤りに対応する確率である．

$\delta = -1 < 0$ (新原料は旧原料よりも悪い) とき d は表示5.1の左の曲線のように分布する．この分布で，棄却域に入らない面積（右側の面積）は，

　　新原料が旧原料よりも悪いにもかかわらず，新原料に切り替える

誤り，すなわち，第2種の誤りの確率 β である．

この対策は，旧原料が新原料よりも良いことが積極的に認められたときにのみ，旧原料を使いつづけることとし，はっきりしないときは新原料に切り替え

ることになる．

αを小さくすると，旧原料が生き残れるチャンスは少なくなる．

対策（2）の意味と性質は，表示5.1のグラフを左右逆にすれば分かるように，第1種の誤りは，

新原料が旧原料に比べ良くないのに，新原料に切り替える誤りである

のに対して，第２種の誤りは，

旧原料が新原料に比べ良くないのに，旧原料を使い続ける誤りである．

この対策は，新しい原料が古い原料よりも良いことが積極的に認められたときにのみ，新しい原料に切り替えることとし，はっきりしないときは古い原料を使い続けるになる．

新しい原料に切り替えるために，多額の設備投資を必要とするような場合には，誤った投資を避けるために，αを小さくすることが要求される．

5.2 第２章 ２項分布に関する推定と検定(1)

演習4(p.29)

表示2.2 の x に，2 と 7 を入力すると，表示5.2の中央に示すように，2 以下の確率が0.0547，7 以下の確率が0.9453 と得られる．

表示 5.2: 2項分布の確率の計算

x	n	π	確率	累積確率	
1	10	0.5	0.0098	0.0107	
8	10	0.5	0.0439	0.9893	
演習4					
2	10	0.5	0.0439	0.0547	0.8906
7	10	0.5	0.1172	0.9453	
演習5					
0	20	0.1	0.1216	0.1216	
0	20	0.15	0.0388	0.0388	
0	20	0.05	0.3585	0.3585	

両者の差を取った 0.8906 が $3 \le x \le 7$ の確率である．これから，2 以下また

は 8 以上を棄却域とするときの α は $1 - 0.8906 = 0.1094$ となる．

演習 5(p.29)

表示5.2 の下に示すように，$x = 0,\ n = 20$ で，$\pi = 0.1, 0.15, 0.05$ を入力すると，$x = 0$ の確率が求められる．

演習 6(p.31)

演習の指示に従って，シート「§1.2」をコピーし，L1 のセルに 0.1, 0.3 を入力すると，表示5.3 の度数表とヒストグラムが得られる．

表示5.3: シミュレーションの結果

π=0.1

個数	度数	ヒストグラム	理論値(%)
0	31	*******************************	34.87
1	46	**	38.74
2	15	***************	19.37
3	8	********	5.74
4	0		1.12
5	0		0.15
6	0		0.01
7	0		0.00
8	0		0.00
9	0		0.00
10	0		0.00
合計	100		100.00

π=0.3

個数	度数	ヒストグラム	理論値(%)
0	2	**	2.82
1	3	***	12.11
2	22	**********************	23.35
3	38	**************************************	26.68
4	22	**********************	20.01
5	8	********	10.29
6	3	***	3.68
7	1	*	0.90
8	1	*	0.14
9	0		0.01
10	0		0.00
合計	100		100.00

得られたヒストグラムの形状は，表示2.4 の対応する折れ線とよく似ている．また，右端に2項分布の確率が計算されている．度数表と比べて見よ．

演習 7(p.44)

表示2.13 の π に 0.30 を，n に 50 を入力する．x の初期値として $n\pi = 50 \times 0.30 = 15$ を入力する．

テキストに述べられている手順に従い，累積確率または上側累積確率が 0.025 以下になる x を求めると，両側検定の棄却域が，9 以下 と 23 以上となる．

表示5.4: 正確法による棄却域の計算

両側検定

x	n	π	確率	累積確率	上側確率
15	50	0.300	0.1223	0.5692	0.5532
9	50	0.300	0.0220	**0.0402**	0.9817
8	50	0.300	0.0110	**0.0183**	0.9927
22	50	0.300	0.0128	0.9877	**0.0251**
23	50	0.300	0.0067	0.9944	**0.0123**

片側検定（下側，上側）

x	n	π	確率	累積確率	上側確率
10	50	0.3000	0.0386	**0.0789**	0.9598
9	50	0.3000	0.0220	**0.0402**	0.9817
20	50	0.3000	0.0370	0.9522	**0.0848**
21	50	0.3000	0.0227	0.9749	**0.0478**

また，累積確率または上側累積確率が 0.05 以下になる x を求めると，片側検定の棄却域が，9 以下，または，21 以上となる．

演習 8(p.46)

表示2.15 に $\pi = 0.3$, $n = 50$, $x = 19$ を入力すると，2項分布の「両側」の欄に 0.2811,「上側」の欄に 0.1406 として，p値が得られる.

表示5.5: 正確法によるp値の計算

			2項分布			正規近似		
π	n	x	下側	上側	両側	下側	上側	両側
0.3	50	19	0.9152	**0.1406**	**0.2811**	0.9175	**0.1400**	**0.2801**

表示5.4 で $x = 19$ を入力しても，同様の結果が得られる.

演習 9(p.47)

表示2.14 の π に 0.30 を，n に 50 を入力する．片側αに 0.025 を入力すると，両側検定の下限が 8，上限が 22 となる.

片側αに 0.05 を入力すると，片側検定の下限が 9，上限が 21 となる.

表示5.6: 正規近似による棄却域の計算

π	n	片側α	期待値	標準偏差	下限	上限	下限	上限
0.3	50	0.025	15.0	3.24	8.15	21.85	8	22
0.3	50	0.05	15.0	3.24	9.17	20.83	9	21

演習 10(p.48)

演習 8 の計算表，表示5.5 の右に解が求められており，両側のp値は 0.2801, 上側のp値は 0.1400 となる.

演習 11(p.48)

この問題は，視聴率が有意に増加したかどうかを知るためであるから，上側に棄却域を設ける 片側検定である.

表示2.15 の計算表に，$\pi = 0.17,\ n = 500,\ x = 96$ を入力すると，表示5.7 に示すように，上側の確率が 0.1069, 0.1056 となる．この値は 0.05 よりも大きいから，視聴率が 17% を超えたということはできない．

表示5.7: πの検定

			2項分布			正規近似		
π	n	x	下側	上側	両側	下側	上側	両側
0.17	500	96	0.9129	**0.1069**	0.2137	0.9145	**0.1056**	0.2113
0.17	500	**99**	0.9557	**0.0561**	0.1122	0.9579	**0.0540**	0.1080
0.17	500	**100**	0.9654	**0.0443**	0.0886	0.9675	**0.0421**	0.0843

x を少しずつ増やしていくと，$x = 99$ で上側確率が 0.0561, 0.0540 であるが，$x = 100$ にすると， 0.0443, 0.0421 といずれも 0.05 以下になる．すなわち，500 人中 100 人が視聴していたとすると，視聴率が17% よりも増加したといえる．

5.3　第3章　2項分布に関する推定と検定(2)

演習12(p.63)

表示3.4 と表示3.6 の計算表で，$n = 100,\ x = 30$ を入力する．π の値を変化させて，下側または上側の確率が 0.05 となる値を求める．

ゴールシークを使って求めた結果を表示5.8 に示す．

信頼率 95% の片側信頼区間を求める問題であるから，確率を0.05 とする．

スコア法による結果は，$\pi > 0.231$ または $\pi < 0.380$ となる．

演習13(p.69)

表示3.10 （正規近似によるの検出力）の$n = 50,\ \pi_0 = 0.5$ をそのままにし，π を 0.3 の前後で変化させる．それに伴い，グラフが変化し，第2種の誤りの

表示5.8: πの区間推定

スコア法

π	n	x	下側	上側
0.380	100	30	0.0500	0.9500
0.231	100	30	0.9500	0.0500

正確法

π	n	x	下側	上側
0.384	100	30	0.0500	0.9684
0.225	100	30	0.9692	0.0500

確率 β の変化する様子を見ることができるであろう．その結果を表示3.9 と比較することにより，表示3.10 の意味の理解を深めることができる．

$\pi_1 = 0.25,\ 0.30,\ 0.35$ に対する β を表示5.9の正規近似の欄に示す．

表示3.9 の $n = 50$ の行から，$\pi_1 = 0.25,\ 0.30,\ 0.35$ に対する検出力を取り出し，表示5.9の検出力の欄に示す．一番右の欄は，1 から 検出力を引いて 求めた β である．

表示5.9: 2つの方法による β の比較

$n = 50,\ \pi_0 = 0.50$

	正規近似法	正確法	
π_1	β	検出力	β
0.35	0.309	0.726	0.274
0.30	0.098	0.915	0.085
0.25	0.015	0.986	0.014

2つの β にはこの程度の差が生じる．

演習 14(p.71)

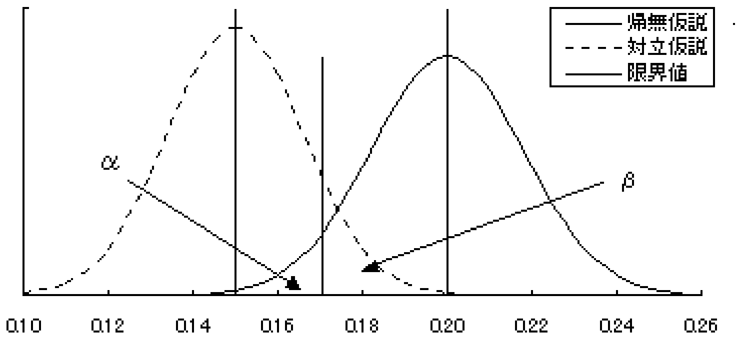
表示3.10で，$\pi_0 = 0.2, \pi_1 = 0.15, \alpha$(片側) $= 0.05$ を入力する．n はとりあえず 300 にする．β は 0.280 となり,「ほぼ確実に」という条件は満たされない．

$\beta = 0.10$ になる n を ゴールシークで求めると,

表示5.10 に示すように，$n = 498$ が得られる．

表示5.10: αとβの関係

n	π0	π1	α(片側)	限界値	β
498	0.2	0.15	0.05	0.171	0.100



すなわち，約 $n = 500$ のサンプルが必要である．

表示5.10 の横軸の目盛りは，このグラフに適するように調整してある．α，βの面積が，それぞれ，5% と 10% になっている．

なお，この値は 表示3.11 の計算表を用いても求めることができる．

演習 15(p.75)

表示3.13 で，旧法の x を 6 に修正すると，p値（片側）が 0.071 となり，この値は 0.05 よりも大きいから，有意に改善されたとはいえない．

5.4　第4章　その他の離散分布

演習 16(p.81)

表示4.3 をコピーし，行数を増やす．
データの本体に第2単元の表示1.13のデータをコピーする．
最後の「相対的外れ」と χ^2検定の部分を表示5.11に示す．

表示5.11: 分割表の χ^2検定の計算結果

相対的外れ	１人１台体制	必要台数のみ	必要台数未達	未導入または検討中
建設	0.39	0.59	-1.49	-1.20
製造	0.14	0.00	-0.15	-0.14
運輸	-0.80	1.23	-1.47	-0.57
販売	**-2.07**	0.88	0.31	1.82
医療	**2.24**	-0.89	-1.45	0.55
ｻｰﾋﾞｽ業	0.49	0.50	-1.59	-0.81
官公庁	**2.01**	**-3.67**	**5.42**	0.49
他	-1.70	0.86	0.80	-0.73

a	8		
b	4	自由度	21
χ^2	81.11	p値	0.0000

$\chi^2 = 81.11$, 自由度 $= 21$ で，p値は 0.0000 となり，高度に有意であることが確かめられた．

「相対的外れ」の表で値が ± 2 外のものは太字（赤）に変えてある．

これから，官公庁は1人1台かまたは未達が多く，必要台数が少ないという特徴がある（第2単元のグラフからも読み取れる）．

演習 17(p.86)

演習16 のヒントに従って，２つの計算表を１つのシートにコピーして，問題の数値を入力すると，表示5.12が得られる．

平均値の差の仮説検定のp値（両側）と分割表のp値はいずれも，0.0368 である．また，前者の $z = -2.088$ を2乗すると後者の $\chi^2 = 4.360$ に一致する．すなわち，どちらの方法も同じ結果を与えることが確認できる．

表示5.12: 2つの割合の差の検定と分割表の χ^2 検定

	n	x	p
放送前	100	29	0.29
放送後	150	63	0.42
合計	250	92	0.37

仮説検定

d	V[d]	se(d)	z	p値(片側)	p値(両側)
-0.13	0.0039	0.0623	-2.088	0.018	0.0368

	知っている	知らない	計
放送前	29	71	100
放送後	63	87	150
計	92	158	250

χ2	4.360
p値	0.0368

p値が 0.05 よりも小さいから，認知率が有意に向上したといえる．

演習18(p.90)

μ, x を入力すると，$= x$, $\leq x$, $\geq x$ の確率の計算表を表示5.13に示す．

表示5.13: ポアッソン分布の確率と累積確率

	B	C	D	E	F
2	μ	x	確率	下側確率	上側確率
3	2	5	**0.0361**	0.9834	0.0527
4	2	6	0.0120	0.9955	**0.0166**
5	2	0	**0.1353**	0.1353	**1.0000**
6	D3:	=POISSON(C3,B3,FALSE)			
7	E3:	=POISSON(C3,B3,TRUE)			
8	F3:	=IF(C3=0,1,1-POISSON(C3-1,B3,TRUE))			

太字が問題の解答である．

（補足）上側の累積確率は，$x-1$ までの下側累積確率を求め，1 から引く．$x=0$ のときは，この計算ができず，常に 1.00 であるから，表示5.13の一番下に示すような関数を用いる．