目　　　次

1. 集めて測る
1.1. 統計学の発見
統計は何をやっているか?　⇒「集めて測る」
集めることで何か見えてくる!

18 世紀ごろから徐々に形成
←　「死亡秩序 mortal order」の発見

現在の歴史人口学のデータでみると
⑴ イングランドの 1811 年の平均寿命推計
農村部 41　10 万未満都市 32　10 万以上 30

⑵ 16-17 世紀イングランドの乳児死亡率
都市:ロンドン 4 教区 0.228　ヨーク 3 教区 0.237
市場町:ゲインズバラ 0.222　バンベリィ 0.165
農村:6 教区平均 0.121 (0.090-0.149)

⑶ 18 世紀イングランドの乳児死亡率
都市:ロンドン 0.319
市場町:ゲインズバラ 0.23　バンベリィ 0.220
農村:11 教区平均 0.165

John Graunt『死亡表 *Bills of Mortality*』1662 年
ちなみに正式の題は
Natural and political observations mentioned in a following index, and made upon the bills of mortality

⑴ 「死には二種類ある!」
「総埋葬数に対して恒常的な比率を保つもの/そうでないもの」
第一種の死:脳卒中、心臓発作、事故、自殺
第二種の死:伝染病(ペスト、しょう紅熱など)

病気による死も集めて測るとちがいがある
→事故や自殺と同じものとちがうもの

事故は偶然によるが集めて測れば法則性
→偶然を支配する法則がある
⇒「極限定理 limit theorem」　個数 N が大きくなると特定の法則にしたがう

⑵「生を超過する死」
都市では死亡が出生をつねに超過する　「都市=蟻地獄」仮説

⑶「変わる数と普遍法則」
男女の出生数がちがう 「出生数の男女比は 1:1 でなく約 14:13(正確には 1.07)」を発見
「アダムとイブ」なら神は 1 対 1 でつくったはず!　なのになぜ???

→誤差と法則性をどう識別するか
⑴の問題でもあるが　出生比≠1 は聖書がらみなので特に衝撃的だったらしい

1.2. 分布の形状を調べる

Graunt の発見「変わる数と普遍法則」

=「出生数は大きく動くのに出生比は恒常的」

少し厳密に定式化してみよう

⇒どうすれば男女の出生数、出生比という数字の集まりをうまく捉えられるか?

1) 集まりが全体としてどのあたりにあるのか

:集まっている位置を示す指標

『平均 mean』　$(1/N)\sum_i(x_i)$

　※以下 $1/N\sum_i(x_i)$で表記する

2) 集まり方がどの程度ばらついているのか

:散らばり方(集まり方)の程度を示す指標

『分散 variance』

$(1/N)\sum_i(x_i-1/N\sum_i(x_i))^2=1/N\sum_i(x_i-m)^2$

慣習的に平均は m、分散は V または s^2 で表す

定義式の意味

① なぜ N で割るか?

　そうしないと具体的な値の数によって大小がちがう

←$\sum_i(x_i)$は x_i の個数に影響されるから N で割って調整する

② なぜ$(x_i-m)^2$ と 2 乗するか?

　そうしないと差の正負が打ち消しあってしまう

集まりは位置と幅でまず押さえる　専門的な言い方では「平均と分散は 1 変数の分布をみる最も基本的な指標」

『分布 distribution』:集まり(散らばり)のあり方

『変数 variable』:統計学がとらえたいと思っている対象　変動する値をもつもの

ロンドンの出生数と出生比に話を戻すと
82 年間の
男性の出生数:平均 5907　分散 2732354
女性の出生数:平均 5533　分散 2531195
男女の出生比:平均 1.07　分散 0.000986
出生数の分散は出生比の分散にくらべて明らかに大きい
「出生数は大きく動くが出生比は恒常的」といえるか?
⇒まだダメ!
元の数の絶対的な大きさが影響する　例えば単位を変えただけで大きさがかわる
その影響をどうすれば外せるか?
→測りたい特性だけをいかに取り出すか?
→直感的な感覚をいかにうまく再現するか?
注目点　⇒平均が大きくなるほど分散は大きい　⇒相殺させればよい!
(1) 分散/平均
ちょっと変……
分散は定義式をみると 2 乗している　だから単位の影響も 2 乗(次元解析の考え方)
ならば
(2) (分散)$^{0.5}$/平均
つまり分散の平方根を平均で割ればよい
分散の平方根を『標準偏差 standard deviation』
とよぶ
『標準偏差』:$(1/N\sum_i(x_i-m)^2)^{0.5}$
標準偏差は s で表す　←"standard deviation"の s
『変動係数 coefficient of variation』
:標準偏差/平均=s/m
82 年間での
男性の出生数:変動係数 0.278
女性の出生数:変動係数 0.286
男女の出生比:変動係数 0.029
たしかに出生比の変動係数は小さい

1.3. 指標から分布をみる

平均と分散以外にも測る指標はありうる

例えば

$1/N\sum_i(x_i-m)^3$

これは正負どちらの方へどの程度ふれているか=歪み具合を示す

→『歪度 skewness』:$(1/N\sum_i(x_i-m)^3)/s^3$

$1/N\sum_i(x_i-m)^4$

これはどの程度真ん中に集まっているか=尖り具合(正確には尖らなさ具合)を示す

→『尖度 kurtosis』:$(1/N\sum_i(x_i-m)^4)/s^4$　または

$(1/N\sum_i(x_i-m)^4)/s^4-3$

「-3」の 3 は正規分布の$(1/N\sum_i(x_i-m)^4)/s^4$

⇒$(1/N\sum_i(x_i-m)^4)/s^4-3$ は正規分布と比べて尖っているかいないかを+-で示す

『平均』　全体的な位置　→x_i の 1 乗に関わる

『分散』　散らばり方　→x_i の 2 乗に関わる

『歪度』　歪み方　→x_i の 3 乗に関わる

『尖度』　尖り方　→x_i の 4 乗に関わる

5 乗項や 6 乗項や……に関わる指標もありえて　それぞれ数の集まり=『分布』の何らかの特徴を示す

経験的には 4 乗項ぐらいで大体わかるので特別な用語はない

しかし理論的に考えると

1 乗項、2 乗項、3 乗項、4 乗項、……は分布のあり方の特徴をそれぞれ表現する

だとすればそれら全部をあつめてくれば分布を一つ特定できるのでは

例えば　こんな感じで

→平均
→分散
分布 →歪み ⇔「分布＝平均＋分散＋…」
→尖り
→……

実は少し工夫すれば本当に数式で成り立つ!

→1 乗項 $1/N\sum_i x_i$
→2 乗項 $1/N\sum_i x_i^2$
分布関数 →3 乗項 $1/N\sum_i x_i^3$
→4 乗項 $1/N\sum_i x_i^4$
→5 乗項 $1/N\sum_i x_i^5$
→……

⇔「分布関数は $1/N\sum_i x_i+1/N\sum_i x_i^2+1/N\sum_i x_i^3+1/N\sum_i x_i^4+\ldots$で捉えられる」

→『母関数 generating function』という発想へ
コンセプトと数式がうまく循環しあう例の一つ

大体の位置は 1 乗項、散らばりは 2 乗項でうまく捉えられる
⇒3 乗項や 4 乗項にも対応する特性がある
⇒各 n 乗項にもそれぞれ対応する特性がある
⇒各 n 乗項を「合わせる」と分布の全体像に

各 step はそれぞれ「集めて測る」ことだが、step ごとにより一般化が進んでいる
step3 や step4 が成立するところが統計学の面白さ step4 が厳密に数式でも成立するところは特にそう
⇒数式を通じて考える数理科学

社会科学では step2 ぐらいまでいくが 3 はめずらしい 4 はほとんどない

- **補足**

散らばり方を示すもう一つの指標

$1/N\sum_i |x_i - m|$　『平均偏差』

男性の出生数:平均 5907　平均偏差 1424

女性の出生数:平均 5533　平均偏差 1372

男女の出生比:平均 1.07　平均偏差 0.0235

男性の平均偏差/平均 0.2411　変動係数 0.278

女性の平均偏差/平均 0.2480　変動係数 0.286

出生比平均偏差/平均 0.0235　変動係数 0.029

ほぼ同じ傾向を示す　あたりまえだが……

⇒どちらも散らばり方の指標としては妥当

1.4. 幅のある判断　確率の導入

Graunt よりもっと厳密に考えてみた人がいた

John Arbuthnot の 1712 年の論文

←I.Newton や J.Swift と同時代人で　イングランド人の仇名"John Bull"の創案者

本当に「1:1」ではないのか?

Arbuthnot の答え方:仮説(H_0)がデータ(D)とどの程度適合するかを見てみよう

←統計的検定の第一歩

H_0:「女性が多い年と男性が多い年は 1/2 の確率でおこる」

D:「男性が多い年が 82 年続いている」

⇒

H_0 の下で D が出現する確率は$(0.5)^{82}$できわめて小さい

だから H_0 であるとはいえない

⇒偶然の誤差を考えた上での判断!

集めて測ることのもう一つの面

もともと集まり=『分布』は幅のあるもの

→その性質も幅のある形で成立していると考えた方が素直

Arbuthnot の判定の結論は「めったに起こらないから 1:1 であるとはいえない」

≠「1:1 ではない」

しかし　ただ「いえない」というだけではなく、*どの程度いえないのかを数値化*してみせる

いえない程度を数値化して考える　程度を測るという発想　⇒幅とか集まりで判断する良さの一つ!

古典的な論理の 0/1 型判断とのちがい

2. 確率変数として見る

2.1. 確率的なモデル

Arbuthnot の判定をもう一度振り返ってみると

H_0「女性が多い年と男性が多い年はそれぞれ 0.5 の確率でおこる」

D「男性が多い年が 82 年続いている」

⇒H_0 の下で D が出現する確率は$(0.5)^{82}$で、きわめて小さいから H_0 であるとはいえない

この考え方をもう一段抽象化すれば

H_0 という性質

　二つの値{女性が多い, 男性が多い}をとり

　各種の確率は P(女性が多い)=0.5、P(男性が多い)=0.5

をもつ変数がある

観察されたデータ D={男性が多い, 男性が多い, 男性が多い, ...}

Arbuthnot の判定:

H_0 の性質をもつ変数から D{男性が多い, 男性が多い, ...}が出現したといえるか?

⇒出現する確率が$(0.5)^{82}$できわめて小さいから、そうであるとはいえない

『確率変数 variable』　=「変動する値を確率的にとるもの」

例えば

　値は{女性が多い, 男性が多い}

　各値をとる確率は P(女性が多い)=0.5、P(男性が多い)=0.5

それぞれ決まった呼び方(定義)があって

『確率変数』X とは

　$\{x_1, x_2, x_3, \ldots\ldots\}$という値を

　それぞれ確率 $P(x=x_i)$でとるもの

$\{x_1, x_2, x_3, \ldots\ldots\}$を『観測値』または『値』

各観測値 x_i に対応するその確率 $P(x=x_i)$をあたえる関数 $f(x_i)$を『確率関数』と呼ぶ

Arbuthnot の判定とは

値域$\{x_1, x_2\}$={男性が多い，女性が多い}

確率関数 $f(x_1)=0.5$、$f(x_2)=0.5$

という確率変数 X があって、観察データ D はその結果であるといえるか

⇒

出現確率は$(0.5)^{82}$だから、いえない

{男性が多い，女性が多い}も数値化すれば

確率変数 X

値域$\{x_1, x_2\}=\{1, 0\}$

確率 $f(x_1)=0.5$、$f(x_2)=0.5$

観測されたデータは{1, 1, 1, 1, 1, ...}

確率変数はふつう大文字で表し　その値をその小文字で表す

例えば「確率変数 X、その値 x」みたいに

なんで値もデジタル化するのか?

←確率関数の特徴をうまく指標化できる

確率変数 X(の確率関数 f(x))はどのような特徴をもつか

① 確率変数 X が出す値は大体どのくらいか

これを『X の期待値』(略して「期待値」)という

$E(X)=\sum x_i \cdot f(x_i)$　←E は期待 expectation の頭文字

Arbuthnot の判定の例では $E(X)=0.5$

② 確率変数 X が出す値はどれだけばらつきがあるか

これを「X の期待値周りのばらつきの期待値」とよぶ(以下「ばらつき期待値(=E(ばらつき))」と省略する)

E(ばらつき)$=\sum(x_i-E(X))^2 f(x_i)$

Arbuthnot の判定の例では

E(ばらつき)

$=(1-0.5)^2\times0.5+(0-0.5)^2\times0.5$

$=0.25\times0.5+0.25\times0.5=0.25$

同じように

③ 歪み期待値

E(歪み)$=\sum(x_i-E(X))^3f(x_i)$

④ 尖り期待値

E(尖り)$=\sum(x_i-E(X))^4f(x_i)$

『分布』の捉え方と発想は同じ

なので①は平均、②は分散とも呼ばれる

ただし　確率変数の期待値やばらつき期待値はその挙動を確率的に示すもので、実際の観察される値(=データ)の平均や分散とは別

期待値やばらつき期待値は

確率関数 f(x)　すなわち

観測値$\{x_1, x_2, x_3, \ldots\ldots\}$に、各値が出現する確率の大きさ$\{P(x_1), P(x_2), P(x_3), \ldots\}$を対応させる関係　の特徴を指標化したもの

(←指標にするために値をデジタル化した)

期待値やばらつき期待値をみれば、どんな範囲で観測値が出てきやすいかはわかるが、実際に観測されるデータの特徴ではない

なので正式には

期待値は『母平均』、ばらつき期待値は『母分散』と、『母』をつけて表記する

例えば Arbuthnot の判定であれば

仮定した確率関数の母平均は 0.5 で母分散 0.25 だから
観測値は平均が 0.5 ぐらいで、分散は 0.25 ぐらいになりやすい(=と期待できる)が
実際には{1, 1, 1, ...}というデータが出ることもありうる

このデータは平均 1 分散 0
つまり　母平均 0.5、母分散 0.25 の確率関数からでも　平均 1 分散 0 というデータは出現しうる

母平均や母分散は確率変数の挙動の特徴を示す値、データの平均や分散は出てきた観測値の分布の特徴を示す値で
対応はあるが　ちがった世界に属する
＊混同するとすごく混乱するので注意!

- **母平均と実際の観測値の関係**

これも数値例で考えるとイメージがしやすい
例えば
ロンドンの性別出生という確率変数 X が
値域$\{x_1, x_2\}=\{1, 0\}$
確率関数 $f(x_1)=0.5, f(x_2)=0.5$
にしたがうとしよう

すでに計算したように　この X の母平均は 0.5、母分散は 0.25

では　この確率変数を 4 回分(=4 年分)観測したら、実際のデータはどんなものになるか?

論理的にありうるすべての場合を書き出すと、
$\{1, 1, 1, 1\}$~$\{0, 0, 0, 0\}$の $2^4=16$ 通り

{1, 1, 1, 1}なら

　　データの平均 1　分散 0

{1, 1, 1, 0}…{0, 1, 1, 1}なら

　　データの平均 0.75　分散 0.1875

{1, 1, 0, 0}…{0, 0, 1, 1}なら

　　データの平均 0.5　分散 0.25

{1, 0, 0, 0}…{0, 0, 0, 1}なら

　　データの平均 0.25　分散 0.1875

{0, 0, 0, 0}なら

　　データの平均 0　分散 0

同じ母平均や母分散の下でも　データではさまざまな平均や分散が出るが
その出やすさがちがう

{1, 1, 1, 1}が出る確率は$(0.5)^4\times1$

→データが平均 1 分散 0 になる確率は$(0.5)^4\times1=0.0625$

{1, 1, 1, 0}…{0, 1, 1, 1}が出る確率は$(0.5)^4\times4$

→データが平均 0.75 分散 0.1875 になる確率は$(0.5)^4\times4=0.25$

{1, 1, 0, 0}…{0, 0, 1, 1}が出る確率は$(0.5)^4\times6$

→データが平均 0.5 分散 0.25 になる確率は$(0.5)^4\times6=0.375$

{1, 0, 0, 0}…{0, 0, 0, 1}が出る確率は$(0.5)^4\times4$

→データが平均 0.25 分散 0.1875 になる確率は 0.25

{0, 0, 0, 0}が出る確率は$(0.5)^4\times1$

→データが平均 0 分散 0 になる確率は 0.0625

つまり
データの平均や分散は必ずしも母平均や母分散と同じ値にはならないが
母平均 0.5 に近いデータの平均、母分散 0.25 に近いデータの分散が出やすい
とはいえる

2.2. 母平均の意味と意義

確率変数とその確率関数にはいろいろある

一番簡単で一番旧く、今もよく使われるものは高校の数学ですでにおなじみ!
「白い球が p、赤い球が q の比率(ただし p+q=1)で入っている袋から、n 回球を取り出してそのうち x 個が白い球である確率を求めよ」

これも確率変数 X とみなせる　図に描くと

	「x の値」
	→0
確率変数 X	→1
=「白い球の数」	→2
	→…
	→n

この確率変数 X は
　値域={x | 0, 1, 2, 3, 4, ……, n}
　確率変数 $f(x)={}_nC_x \cdot (p)^x \cdot (1-p)^{n-x}$　(x は 0 から n までの整数)

この形の確率関数を「二項分布」「ベルヌーイ分布」とよび、確率変数を「二項分布にしたがう確率変数」とよぶ
二項分布の特性
　母平均 np、母分散 np(1-p)

確率変数の母平均や母分散は「その変数の基本性能」　⇔観測(観察)された値の集まりとは必ずしも一致しない
　例えば　D もありうる
「十分多い観測」という条件の下で、観測された値の平均や分散は、母平均や母分散と一致する

母平均や母分散は独自の演算規則

母平均すなわち期待値 $E(\cdot)$では

1) $E(c)=c$ (c は定数 constant を表す)

2) $E(X+c)=E(X)+c$

3) $E(cX)=cE(X)$

4) $E(X+Y)=E(X)+E(Y)$

母分散は通常 $V(\cdot)$で表すが、$V(\cdot)$では

1) $V(c)=0$

2) $V(X+c)=V(X)$

3) $V(cX)=c^2V(X)$

4) $V(X+Y)=V(X)+V(Y)+2E\{(X-E(X))(Y-E(Y))\}$

つまり加法則は$(X-E(X))(Y-E(Y))$の期待値(=母平均)が 0 になるとき以外は成立しない

$E\{(X-E(X))(Y-E(Y))\}$を、X と Y の「共分散 covariance」といい、$Cov(X, Y)$と表記する

- **確率変数同士の演算の考え方**

確率変数 X と Y の和 X+Y とは (X と Y 各々の値 x_i と y_j とすると)

「ありうる全ての値の組み合わせ(x_i, y_j)に関して、x_i+y_j を値とし、x_i かつ y_j となる(=x_i と y_j が同時に成立する)確率 $P(x_i, y_j)$を値 x_i+y_j の出現確率とする確率変数」

※値域が x_i+y_j で定義され、各「x_i+y_j」に対してその確率が定義できている(=確率関数があたえられている)ので、X+Y も確率変数

同じように

確率変数の積 XY は「ありうる全ての値の組み合わせ(x_i, y_j)に関して、x_iy_j を値とし、x_i かつ y_j となる(=x_i と y_j が同時に成立する)確率 $P(x_i, y_j)$を値 x_iy_j の出現確率とする確率変数」

になる

$E(\cdot)$の加法則を一番簡単な場合で証明すると

X は x_1, x_2、Y は y_1, y_2 という 2 つの値をもつとする　$E(\cdot)$は確率変数の値にその出現確率をかけたものだから

$$
\begin{aligned}
&E(X+Y)\\
&=(x_1+y_1)P(x_1, y_1)+(x_1+y_2)P(x_1, y_2)+(x_2+y_1)P(x_2, y_1)+(x_2+y_2)P(x_2, y_2)\\
&=x_1P(x_1, y_1)+y_1P(x_1, y_1)+x_1P(x_1, y_2)+y_2P(x_1, y_2)\\
&\quad+x_2P(x_2, y_1)+y_1P(x_2, y_1)+x_2P(x_2, y_2)+y_2P(x_2, y_2)\\
&=x_1P(x_1, y_1)+x_1P(x_1, y_2)+x_2P(x_2, y_1)+x_2P(x_2, y_2)\\
&\quad+y_1P(x_1, y_1)+y_1P(x_2, y_1)+y_2P(x_1, y_2)+y_2P(x_2, y_2)\\
&=x_1\{P(x_1, y_1)+P(x_1, y_2)\}+x_2\{P(x_2, y_1)+P(x_2, y_2)\}\\
&\quad+y_1\{P(x_1, y_1)+P(x_2, y_1)\}+y_2\{P(x_1, y_2)+P(x_2, y_2)\}
\end{aligned}
$$

$P(x_1, y_1)+P(x_1, y_2)$は x_1 かつ y_1 である確率と x_1 かつ y_2 である確率の和だから、

$P(x_1, y_1)+P(x_1, y_2)=P(x_1)$

したがって、

$=x_1P(x_1)+x_2P(x_2)+y_1P(y_1)+y_2P(y_2)$

$=E(X)+E(Y)$

で $E(\cdot)$では加法則がなりたつ

$V(\cdot)$についてみると

V(X)は確率変数 X の各値 x_i から母平均 E(X)を引いたものの 2 乗、すなわち$(x_i\text{-}E(X))^2$ に x_i の出現確率 $P(x_i)$をかけたものだから、$(x_i\text{-}E(X))^2$ を値にする確率変数「$(X\text{-}E(X))^2$」の母平均に等しい

同様に、V(X+Y)は　$(x_i+y_j\text{-}E(X+Y))^2$ に x_i+y_j の出現確率 $P(x_i, y_j)$をかけたものだから、$(x_i+y_j\text{-}E(X+Y))^2$ を値にする確率変数「$(X+Y\text{-}E(X+Y))^2$」の母平均に等しい

※X と Y は確率変数、E(X+Y)は $E(\cdot)$の加法則より E(X)+E(Y)で E(X)も E(Y)も特定の値すなわち定数だから E(X+Y)も定数、それゆえ X+Y-E(X+Y)は確率変数であり、その 2 乗$(X+Y\text{-}E(X+Y))^2$ も確率変数であり、各値$(x_i+y_j\text{-}E(X+Y))^2$ の確率は $P(x_i, y_j)$であたえられる

したがって、V(X+Y)は

$$
\begin{aligned}
&V(X+Y)\\
&=\textstyle\sum_{i,j}(x_i+y_j-E(X+Y))^2 f(x_i, y_j)\\
&=\textstyle\sum_{i,j}(x_i+y_j-E(X+Y))^2 P(x_i, y_j)\\
&=E(X+Y-E(X+Y))^2\\
&=E\{X^2+Y^2+(E(X+Y))^2+2XY-2XE(X+Y)-2YE(X+Y)\}\\
&=E\{X^2+Y^2+(E(X)+E(Y))^2+2XY-2X(E(X)+E(Y))-2Y(E(X)+E(Y))\}\\
&=E\{X^2+Y^2+(E(X))^2+2E(X)E(Y)+(E(Y))^2+2XY-2XE(X)-2XE(Y)-2YE(X)-2YE(Y)\}\\
&=E\{X^2-2XE(X)+(E(X))^2+Y^2-2YE(Y)+(E(Y))^2+2XY+2E(X)E(Y)-2XE(Y)-2YE(X)\}\\
&=E\{X^2-2XE(X)+(E(X))^2+Y^2-2YE(Y)+(E(Y))^2+2(X-E(X))(Y-E(Y))\}\\
&=E\{X^2-2XE(X)+(E(X))^2\}+E\{Y^2-2YE(Y)+(E(Y))^2\}+E\{2(X-E(X))(Y-E(Y))\}\\
&=E\{(X-E(X))^2\}+E\{(Y-E(Y))^2\}+2E\{(X-E(X))(Y-E(Y))\}\\
&=V(X)+V(Y)+2E\{(X-E(X))(Y-E(Y))\}\\
&=V(X)+V(Y)+2Cov(X, Y)
\end{aligned}
$$

Cov(X, Y)が正になる場合とは
→x が期待値 E(X)より大きく(小さく)なるときは、y もその期待値 E(Y)より大きく(小さく)なる関係が
Cov(X, Y)が負になる場合とは
→x が期待値 E(X)より小さく(大きく)なるときは、y はその期待値 E(Y)より大きく(小さく)なる関係が

それぞれある場合にあたる

株式でいえば

株 X が暴騰するときは株 Y も暴騰し、X が暴落するときは Y も暴落するような関係にあれば Cov(X, Y)は正になる
$\Rightarrow V(X+Y)>V(X)+V(Y)$
株 X が暴騰するときは株 Y は暴落し、X が暴落するときは Y は暴騰するような関係にあれば Cov(X, Y)は負になる
$\Rightarrow V(X+Y)<V(X)+V(Y)$

2.3. モーメントと分布指標

もうちょっと一般化してみよう……

確率変数(確率関数)の挙動を示す値を「理論値(母数) parameter」とよぶ

　E(·)や V(·)はその一つ

※確率変数 X に対して E(X)や V(X)は特定の値をとる　E(·)や V(·)は確率変数を特定の値に変換する演算でもある

E(·)と V(·)だけがわかればその確率関数の挙動は完全にわかるのか?

⇒一般にはダメ　全くちがう確率関数がたまたま同じ E(·)や V(·)をもちうるから

では　ある確率変数の確率関数の挙動を完全に特定できる指標はあるのか?

ある　⇒『モーメント』

確率変数 X における、その値の n 乗 x^n を値とする確率変数 X^n の期待値=$E(X^n)$を　『X の(原点周りの)n 次のモーメント』と呼ぶ

各次のモーメントは次のような性質をもつ

① E(X)すなわち母平均は 1 次のモーメント

② 母分散 $V(X)=E\{(X-E(X))^2\}$は

$E\{(X-E(X))^2\}$

$=E\{X^2-2XE(X)+(E(X))^2\}$

$=E(X^2)-E(2XE(X))+E\{(E(X))^2\}$

$=E(X^2)-2E(X)E(X)+(E(X))^2$

$=E(X^2)-2E(X)^2+(E(X))^2$

$=E(X^2)-(E(X))^2$

　つまり V(X)は 2 次のモーメント $E(X^2)$から 1 次のモーメント E(X)の 2 乗をひいたもの

同様に

③ 歪み期待値 $E\{(X-E(X))^3\}$
$=E(X^3)-3E(X)E(X^2)+2(E(X))^3$
④ 尖り期待値 $E\{(X-E(X))^4\}$
$=E(X^4)-4E(X)E(X^3)+6(E(X))^2E(X^2)-3(E(X))^4$
モーメントは特に『モーメント母関数 moment generating function』で重要になる

モーメント母関数とは各次のモーメントを係数にもつ多項式
$1+E(X)t+\{E(X^2)/(2!)\}t^2+\{E(X^3)/(3!)\}t^3+\ldots$
のこと
この t(と X)の関数を n 回微分して t=0 とすれば n 次のモーメントになる　期待値や母分散も計算できる

よけいややこしい気がするが(^^
$1+E(X)t+\{E(X^2)/(2!)\}t^2+\{E(X^3)/(3!)\}t^3+\ldots$
$=E(e^{tX})$
　という簡単な確率関数 e^{tX}(e は自然対数の底)の期待値の形に変形できるので、元の確率関数 f(x)よりあつかいやすいことが多い
　例えば二項分布 ${}_nC_x\cdot(p)^x\cdot(1-p)^{n-x}$ のモーメント母関数は$(e^tp+q)^n$

他にもいろいろ便利な特性があるので　確率変数の特徴を捉える上で強力な道具になる
例えば
1) 元の確率関数 f(x)と一対一に対応する
　モーメント母関数が$(e^tp+q)^n$ なら二項分布 ${}_nC_x\cdot(p)^x\cdot(1-p)^{n-x}$ だ、とか
2) 2 つの確率変数 X と Y が独立なら、その和 X+Y は確率変数で、そのモーメント母関数は X と Y 各々のモーメント母関数の積になる

でこの形、どこかで見たような気がしないか?
1.3 で述べた話……

　　　　→平均
　　　　→分散
分布　　→歪み　　⇔「分布＝平均＋分散＋…」
　　　　→尖り
　　　　→……

という考え方は少し工夫すれば本当に数式で成り立つ!

　　　　　→1 乗項 $1/N\sum_i x_i$
　　　　　→2 乗項 $1/N\sum_i x_i^2$
分布関数　→3 乗項 $1/N\sum_i x_i^3$
　　　　　→4 乗項 $1/N\sum_i x_i^4$
　　　　　→5 乗項 $1/N\sum_i x_i^5$
　　　　　→……

⇔「分布関数は $1/N\sum_i x_i+1/N\sum_i x_i^2+1/N\sum_i x_i^3+\ldots$ で捉えられる」
この「$1/N\sum_i x_i+1/N\sum_i x_i^2+1/N\sum_i x_i^3+\ldots$」にあたるのが　モーメント母関数
$1+E(X)t+\{E(X^2)/(2!)\}t^2+\{E(X^3)/(3!)\}t^3+\ldots$

つまり各次のモーメントがわかれば
⇒モーメント母関数がわかり
⇒元の確率関数がわかる

これ自体は確率変数の世界での話だが
「*「十分多い観測」という条件の下で*、観測された値の平均や分散は、母平均や母分散と一致する」ので

十分多い観測の下では、得られたデータの n 乗項 $1/N\sum_i x_i$(したがってデータの平均や分散や歪みや尖りや……)から
どんな確率関数を持つ確率変数かが推定できる
⇒『モーメント法』

2.4. 数理統計学のコンセプト

確率変数という考え方を使った判断の方法が『数理統計学(推測統計学)』

天下り的&大雑把にいえば、数理統計学とは

a) 観察されたデータに対して特定の確率変数をモデルとしてあてはめることでデータの出てくるしくみを推測する
b) ある仮説を特定の確率変数によってモデル化することで、その仮説が観察されたデータにどのくらいあてはまるかを(数量的に)判定する

a)を『推定』、b)を『検定』とよぶ
例えば Arbuthnot の判定は歴史上最初の検定の例にあたる

もう一つ重要なポイント!
数理統計学では 2 種類の数値が出てくる
1) 観察されたデータ上での数値
2) 確率変数の挙動を示す数値
1)を『観測値』、2)を『理論値(母数) parameter』とよぶ

2.1 で述べたように　この二つをきちんと区別できることが大切
二つはちがった種類の概念だが、*十分多い観察の下では*、それぞれ対応する値がある

例えば
$E(\cdot)$:理論値の一つで、十分多い観察の下では観測値の平均 m と一致してくる
$V(\cdot)$:理論値の一つで、十分多い観察の下では観測値の分散 s^2 と一致してくる

理論値(母数)はふつうギリシア文字で表す
　日本語では「母」をつける

例えば
$E(\cdot)$はμ(みゅー)　←平均 mean の頭文字"m"にあたるギリシア文字　=「母平均」

$V(\cdot)$はσ^2(しぐま 2 乗)　←分散 s^2 の平方根である標準偏差 standard deviation の頭文字"s"にあたるギリシア文字　=「母分散」

モーメント $E(X^n)$は $\mu_1, \mu_2, \ldots\ldots, \mu_n$
(本当は平均 m が最初に使われたときは一次のモーメント moment の m だった)

けっこう混同してしまうので注意が必要
混乱しやすい有名な例

中心極限定理:「平均 μ、分散 σ^2 の確率変数 Z から出現する N 個の観測値の平均 $1/N\sum z_i$ は、漸近的に平均 μ、分散 σ^2/N の正規分布にしたがう」

……??だが　丁寧に書き直すと
中心極限定理:「<u>母平均 $E(Z)=\mu$、母分散 $V(Z)=\sigma^2$</u> の確率変数 Z から出現する N 個の観測値の平均 $1/N\sum z_i$ は、<u>それ自体がある確率変数の観測値になり</u>、N が無限大に近づくにつれて、<u>その出現の仕方は母平均 $E(1/N\sum z_i)=\mu$、母分散 $V(1/N\sum z_i)=\sigma^2/N$ の正規分布</u>にしたがう」

確率変数の観測値の平均は、それ自体が確率変数の値になり、<u>元の確率変数の確率関数の如何にかかわらず</u>正規分布にしたがう!

3. 確率変数の応用

3.1. 連続型確率変数

2.1 で定式化した確率変数は正確には確率変数の 1 つの型
⇒『離散型確率変数』　とよばれる

確率変数には 2 つの種類がある
離散型 discrete と連続型 continuous

離散型は「ある確率で特定の値をとる確率変数」
連続型は「ある確率である幅に値が入る確率変数」　いわば幅でおさえる捉え方

ロンドンの出生数でいえば
「ロンドンの女性出生力」という確率変数として
a）離散型で
　$\{x_1, x_2, x_3, \ldots\}=\{4688, 4457, 4102, 4590, \ldots\}$
　確率関数 $f(x_i)=0.012$
という変数も想定できるが、
b）連続型で
　~3000 が 0.073、~3500 が 0.098、~4000 が 0.037 という確率で出てくる
という変数も想定できる

実際に観察される数値はつねに特定の値=離散だが　確率変数という捉え方自体が抽象化=モデル化したもの
実現値が特定の値だからといって一つ一つに確率が決まっていると考える必要はない

より便利に使える方でよい　ま大体ってこと!

例えば
1）区間でみていくとわかることも
ロンドンの女性出生数と男性出生数も区間できって見てみると……
500 ずれているがグラフの形状はよく似ている

→区間で切ることで規則性が見えることもある

2）連続型の方が数学的にあつかいやすい
実際にはこれが大きい　最初は離散型の近似計算法だった
←コンピュータができるまでは積分で計算する方がずっと簡単だった

連続型確率変数 X とは
　$-\infty<x<+\infty$のなかで値をとり
　幅 $a<x<b$ になる確率 $P(a<x<b)$が決まっている「何か」=変数である
＊「$<$」が厳密にいえるかはおいておく(^^

このとき $P(a<x<b)$をあたえる関数 $f(x)$
　$P(a<x<b)=\int_a^b f(x)dx$　ただし$\int f(x)dx=1$(特に限定しない場合$\int$は$-\infty\sim+\infty$)
を『確率密度関数』とよぶ

これは離散型の確率関数にあたるもの
　和記号$\sum\sim$⇔積分記号$\int\sim dx$
離散型なら
$P(a<x<b)=\sum_{xi=a}^{b} f(x_i)$　ただし$\sum_i f(x_i)=1$
連続型なら
$P(a<x<b)=\int_a^b f(x)dx$　ただし$\int f(x)dx=1$
つまり$\sum\sim$を$\int\sim dx$ に置き換えたもの

連続型でも離散型と同じく　母平均(=期待値)や母分散などの理論値が定義されている

① 母平均 $E(X)=\int xf(x)dx$　(特に指定なければ積分∫は-∞から+∞まで)
② 母分散 $V(X)=E(X-E(X))^2=\int(x-E(X))^2f(x)dx$
これで数学的にも自然な定義になる

連続型の確率密度関数は指数関数などの形をとることが多い

連続型で最も有名なのが
『正規分布 normal distribution』

正規分布とは　その確率密度関数が
$f(x)=1/(2\pi\beta^2)^{0.5}\cdot\exp_e\{-(x-\alpha)^2/(2\beta^2)\}$
e は自然対数の底、$\exp_e$~は e の~乗
という形になるもの

母平均は $E(\cdot)=\int xf(x)dx=\alpha$
母分散は $V(\cdot)=\int(x-\mu)^2f(x)dx=\beta^2$

正規分布ではその挙動を示す 2 つの母数(理論値)α と β の値は母平均と母分散にちょうどひとしい

ので　慣例的に
$f(x)=1/(2\pi\sigma^2)^{0.5}\cdot\exp_e\{-(x-\mu)^2/(2\sigma^2)\}$
と表記するが　論理的には?
2 つの母数がたまたま母平均と母分散に一致する　と考えた方がよい

母平均と母分散は直接計算できるが　正規分布のモーメント母関数が $\exp_e\{\alpha t+(\beta^2t^2)/2\}$になることから簡単に求まる

正規分布も確率密度関数よりもモーメント母関数が簡単であるものの 1 つ

3.2. 二項分布を使ったモデル
確率変数というモデルを使った判断
を実例で説明しよう

まずモデルに使う確率(密度)関数から
いろいろある
例えば正規分布　連続型の確率密度関数で最も旧くまたよく使われる

離散型の確率関数で最も旧くまたよく使われるのが『二項分布』
$f(x)={}_nC_x \cdot (p)^x \cdot (1-p)^{n-x}$　(x は 0 から n までの整数)
二項分布は「あることが生起する確率が p の場合に、n 回やってみてそのうち x 回それが実際に起こる確率」にあたる

母比率 p は理論値なのでふつうギリシア文字を使うが、円周率 π と区別するため p を使う

ロンドンの出生比の問題
これも二項分布をモデルに使える

Arbuthnot の判定
1) H_0「男性が多い年と女性が多い年はそれぞれ 0.5 の確率でおこる」
2) 実際のデータ D は「男性が多い年が 82 年続いている」
3) H_0 が正しければ D が出る確率は$(0.5)^{82}$　これはきわめて小さいから H_0 だとはいえない

これも二項分布
→「男性が多い年が生起する確率が 0.5 の場合に、82 回やってみてそのうち 82 回男性が多い年になる確率」
を計算している

図に描けば

　　　　　→男性が多い年 or 女性が多い年

　　　　　→男性が多い年 or 女性が多い年

「袋」　　→男性が多い年 or 女性が多い年

　　　　　→……

　を　82 回やって「男性が多い年」が 82 個

⇒「白い球と赤い球が 0.5 ずつの袋から、82 回球を取り出して 82 個白い球である確率」

と同じ

実際　二項分布の確率関数に n=82、x=82 を代入すれば

$f(x)={}_{82}C_{82}\cdot(0.5)^{82}\cdot(1-0.5)^{0}$

$=(0.5)^{82}$

このモデルはどれだけ適切だろうか

疑問点が 2 つある

a）出生は本来個体レベルでおこる

→「1:1」は男性が多い年の出現確率というより、出生児一人一人の男女の出現確率

　実際、男性の多い年は 14/27 で出現するわけではない

b）男性がどのくらい多いかがわかっている

→男性の比率は 0.5027~0.5362 つまり「男性の比率が 0.5027~0.5362 である」年が 82 年間つづいたという情報を取り込んだ方がより良い

a)と b)をうまく取り込んだモデルをつくれないだろうか?

ここで

「出生児が男性である確率が 0.5 である場合に、1 年の出生児全体での男性の比率が 0.5027~0.5362 である」確率が求まれば、それが 82 年間つづいた確率はその 82 乗

したがって「出生児が男性である確率が 0.5 である場合に、1 年の出生児全体での男性の比率が 0.5027~0.5362 である」確率が計算できればよい

これには 2 つやり方がある

★　やり方 1

二項分布で直接計算する

82 年間のロンドンの平均出生児数を 11429 とすると

「出生児が男性である確率が 0.5 である場合に、1 年の出生児全体での男性の比率が 0.5027~0.5362 である」=「出生児が男性である確率が 0.5 である場合に、11429 人の出生児のなかで男性が 5745 から 6128 である確率」

→出生児は男性 or 女性
→出生児は男性 or 女性
「袋」　→出生児は男性 or 女性
→出生児は男性 or 女性
→……

で 11429 回やって「出生児は男性」が 5745~6128 個

⇒「白い球と赤い球が 0.5 ずつの袋から、11429 回球を取り出して、白い球が 5745 個から 6128 個である確率」

これは一つずつのケースを足していけばよい

例えば

5745 個である確率は

$_{11429}C_{5745}\cdot(0.5)^{5745}\cdot(0.5)^{11429-5745}$

同様にして　5746 個である確率は

$_{11429}C_{5746}\cdot(0.5)^{5746}\cdot(0.5)^{11429-5746}$

5747 個である確率は

$_{11429}C_{5747}\cdot(0.5)^{5747}\cdot(0.5)^{11429-5747}$

……

6128 個である確率は

$_{11429}C_{6128}\cdot(0.5)^{6128}\cdot(0.5)^{11429-6128}$

まとめて書けば

$\sum_{x=5745}^{6128}\ {}_{11429}C_{x}\cdot(0.5)^{x}\cdot(0.5)^{11429-x}$

実際にこれを計算した人がいる!

s'Gravesande の論文「日々の男女の出生率から引出されることを例にとり、この世に生起する事象は神の支配によることの注意深い数学的証明」（おーい……)

結果だけ紹介する　平均出生児数を 11429 として

$\sum_{x=5745}^{6128} {}_{11429}C_x \cdot (0.5)^x \cdot (0.5)^{11429-x}$

=0.292

　その 82 乗は$(0.292)^{82}$

Arbuthnot の判定では$(0.5)^{82}$ だから、それよりさらに小さい値になる

←より正確にモデル化することで　もっとありえないといえる

★　やり方 2

近似計算する

「出生児が男性である確率が 0.5 である場合に、1 年の出生児全体での男性の比率が 0.5027~0.5362 である」

1 回 1 回の出生は

「袋」　→出生児は男性 or 女性

男性を「1」、女性を「0」の値に置き換えると

「袋」　→1or0

これも二項分布にしたがう確率変数

変数 Z　値域{男性，女性}={1, 0}

　　　　確率関数 $f(z)={}_1C_z \cdot (0.5)^z \cdot (0.5)^{1-z}$

p=0.5 の袋のなかから n=1 回球を取り出して、それが白(=男性=1)か赤(=女性=0)か

Z は二項分布にしたがうので　その母平均と母分散は

　$E(Z)=np=0.5$

　$V(Z)=np(1-p)=0.25$

この Z の値を 11429 回観察して、平均 $1/11429\sum z_i$ をとったのが「1 年の出生児数 11429 での男性の比率」

$1/11429\sum z_i$ は確率変数の観測値の平均　→

中心極限定理から、これも確率変数の観測値とみなすことができ、その出方は観察回数 z_i が大きければ近似的に正規分布にしたがう

中心極限定理:「母平均 $E(\cdot)=\mu$、母分散 $V(\cdot)=\sigma^2$ の確率変数 Z からの N 個のサンプルの平均 $1/N\sum z_i$ は、漸近的に $E(\cdot)=\mu$、$V(\cdot)=\sigma^2/N$ の正規分布にしたがう」!

※この N と二項分布の n を混同しないように

したがって $1/11429\sum z_i$ は

$E(\cdot)=0.5$、$V(\cdot)=0.25/11429=(0.004677)^2$ の正規分布に近似的にしたがう

だから

こういう E と V をもつ正規分布の値が 0.5027~0.5362 の区間に入る確率を求めてやれば

それが $1/11429\sum z_i$ が 0.5027~0.5362 の値をとる確率にほぼ等しい

正規分布に関しては値がどの区間にどのくらいの確率で入るかはすでに計算されている

⇒「正規分布表」!

　『統計学入門』なら p.280

● 正規分布表の読み方

正規分布表は
「正規分布にしたがう確率変数で、母平均より正または負の方向に u 分以上はずれた値が出る確率の大きさ」　を示す

u とは(値-E(·))/(V(·))$^{0.5}$=(値-μ)/σ
つまり　「母平均 μ からその値がどれだけ離れているかを、母標準偏差 σ を 1 単位にして表したもの」

なぜこんな形で表すのか?
正規分布は母平均や母分散の大きさによって特定の幅の値が出る確率がちがう
⇒
それを母平均 0 母分散 1 の正規分布に標準化するため

そうすることで
<u>一枚の表で</u>　<u>全ての正規分布に関して</u>
「母平均より(正または負の方向に)ある値以上はずれた値が出る確率の大きさ」を計算できる!

やり方その 2 に戻っていえば
知りたいのは
「E(·)=0.5、V(·)=0.25/11429=(0.004677)^2 の正規分布にしたがう確率変数が 0.5027~0.5362 の区間に入る確率」　だから
この確率変数で　「0.5027 より大きな値が出る確率の大きさ」から「0.5362 より大きな値が出る確率」を引けば
「0.5027~0.5362 の区間に入る確率」が求まる

例えば「0.5027」であれば
u=(0.5027-0.5)/0.004677=0.577
これは
「母平均 0.5 母分散$(0.004677)^2$の正規分布で 0.5027~の値が出る確率の大きさは、母平均 0 母分散 1 の正規分布で 0.577~の値が出る確率の大きさに等しい」ことを意味する

この大きさは正規分布表の 0.5 の行の、0.07 の列と 0.08 の列との間

をみればよい
ほぼ 0.282
→つまり 0.5027 より大きな値が出る確率は 0.282

「0.5362」であれば
u=(0.5362-0.5)/0.004677=6.97 だから
表に出てこないくらい小さい(表は 4.99 まで)

ほぼ 0
→つまり 0.5362 より大きな値が出る確率は 0

したがって
「出生児が男性である確率が 0.5 である場合に、1 年の出生児全体での男性の比率が 0.5027~0.5362 である」確率は
0.282-0=0.282

やり方その 1 では 0.292 だったから
かなり近い値になっている

- **出生比 14/27 という仮説の妥当性**

同じように
「出生児が男性である確率が 14/27≒0.52 である場合に、17 世紀後半のロンドンぐらいの大きさの集団で、1 年の出生児中の男性の比率が 0.5027~0.5362 である」
確率も計算できる

やり方 2 でやると、
まず出生比の確率変数の E(·)と V(·)は
　$E(Z)=np=0.52$
　$V(Z)=np(1-p)=0.2496$
になる
したがって、$1/N\sum z_i$ は N が十分大きい場合には中心極限定理により
　$E(1/N\sum z_i)=0.52$
　$V(1/N\sum z_i)=0.2496/11429=(0.004673)^2$

の正規分布にしたがう

あとは正規分布表をみればよい

「0.5027 より大きな値が出てくる確率」は
$u=(0.5027-0.52)/0.004673=-3.7$ だから　ほぼ 1
「0.5362 より大きな値が出てくる確率」は
$u=(0.5362-0.52)/0.004673=3.47$ だから　ほぼ 0
したがって「0.5027~0.5362」の値が出てくる確率は　ほぼ 1-0=1
したがって　その 82 乗は　ほぼ 1

だから p=14/27≒0.52 という仮説はデータにあてはまらないとはいえない

3.3. モデルの使い方(お浚い)

重要な点は 2 つ

⑴ あてはめられるモデルは複数ある

例えばロンドンの出生比でいえば

① Arbuthnot のモデルもあれば

② s'Gravesande のモデルもあり

③ s'Gravesande のモデルを連続型でさらに近似したモデルもある

連続型でも離散型でもいい

⑵ そのなかで

a) 仮説の理論によりあったもの

b) データの情報をより活かせるもの

c) 数値計算が楽なもの

を選べば仮説の妥当性についてより適切な評価ができる　あとは

d) 汎用性が高いもの　　も応用では重要

やり方その 2 は a)~d)をみたす

⇒現在の統計的検定の原型　つまり

<u>ある仮説</u>:*「出生児が男性の確率 p が 0.5」あるいは「p が 0.52」*

が正しければ

<u>観測値から求められる量</u>:$1/N\sum z_i$

が

<u>特定の確率分布</u>:$E(\cdot)=np$、$V(\cdot)=\{np(1-p)\}/N$ *の正規分布*

にしたがう確率変数になる

ことを使って　その量 $1/N\sum z_i$ の出現確率の大きさで仮説「p が 0.5」や「p が 0.52」の正しさを評価する手続き

この「観測値から求められる量」を『検定統計量 test statistic』とよぶ(仮説の検証 test に使われる統計的な数値 statistic という意味)

◎試験では検定は必ずでるので

<u>正規分布と t 分布とカイ 2 乗分布</u>

<u>の分布表</u>

を用意しておくこと!

4. 統計的推定

4.1. 視聴率のしくみ

数理統計学(推測統計学):

『確率変数 variable』　=「変動する値を確率的にとるもの」を使ったモデルで考えていく営み

いろいろ例があるが　身近な一つが「視聴率」

視聴率にもいろいろある

a）個人視聴率と世帯視聴率

b）瞬間視聴率と番組視聴率

測定法はいくつかありそれぞれの妥当性も重要な問題だが今は無視する

以下では 1 世帯に 1 台 TV があると仮定する

いわゆる「視聴率」とはサンプル調査の視聴率=サンプル集団での「視ている」比率

『サンプル調査』:本来調査したい対象から一定のやり方で一部を抜き出して調べる

　抜き出したものが『サンプル(標本)』

　本来調査したい対象が『母集団』

⇒母平均とか母分散の「母」と同じ意味

視聴率調査でのサンプルの大きさ(『サンプル規模 sample size』)は

　関東·関西·名古屋で 600 世帯

　札幌·仙台·広島·福岡などが 200 世帯

母集団の大きさは例えば関東地区全体の世帯数は 1455 万世帯

少し小さすぎる?　サンプル調査で視聴率 30%や 20%が出てもどのくらいあてになる?

サンプル視聴率は母集団の視聴率をある程度反映するだろうが　本当の視聴率(=関東地区全体の視聴率)と同じになる保証はない　むしろちがうのがあたりまえ

<u>重要なのはその精度が計算できる=数値的に評価できる!</u>

4.2. サンプル調査の精度

注目する点　⇒サンプル視聴率の値は二項分布にしたがう確率変数　になる

つまり
「関東地区で母視聴率(母比率)が p である場合に、n 個の世帯取り出してそのうち x 個が視ている確率は?」

　　　　→番組を視ている or 視ていない
　　　　→番組を視ている or 視ていない
『袋』　→番組を視ている or 視ていない
　　　　→番組を視ている or 視ていない
　　　　→……

で
n 回取り出してそのうち x 回「番組を視ている」が出る確率を求めるのと同じ

ロンドンの出生比(Arbuthnot の判定)の話では

　　　　→男の子 or 女の子
　　　　→男の子 or 女の子
『袋』　→男の子 or 女の子
　　　　→男の子 or 女の子
　　　　→……

で
「出生児は男」の確率が 0.5 や 0.52 である場合に n 回取り出して x 回「男の子」が出る確率を求めた

それが
「男の子」　→「番組を視ている」
「出生比 0.5」　→「母視聴率」
に変わるだけ

二項分布の確率関数

$f(x)={}_nC_x \cdot p^x \cdot (1-p)^{n-x}$

を使えば

サンプル規模(n の大きさ)ごとに　サンプル視聴率 x/n の値が出現する確率が計算できる

n=1 すなわち 1 人だけをサンプル集団として調査する場合

サンプル視聴率 x/n は「0%」「100%」どちらかで

「0%」　${}_1C_0 \cdot p^0 \cdot (1-p)^{1-0}$

「100%」　${}_1C_1 \cdot p^1 \cdot (1-p)^{1-1}$

n=2 だと x/n は「0%」「50%」「100%」のどれかで

「0%」　${}_2C_0 \cdot p^0 \cdot (1-p)^2$

「50%」　${}_2C_1 \cdot p^1 \cdot (1-p)^1$

「100%」　${}_2C_2 \cdot p^2 \cdot (1-p)^0$

→p があたえられればサンプル視聴率 x/n の値の出方は求められる

=母視聴率が p のときにサンプル視聴率 x/n としてどの値が出やすいかを計算できる

サンプル調査の特性

1) サンプル視聴率 x/n はどんな人がサンプルになるかによって変わる確率変数

2) サンプル視聴率としてどの値がどのくらい出やすいかはサンプル規模 n と母視聴率 p だけで決まる(二項分布だから当たり前だが)

→<u>母集団の大きさは関係ない</u>!

3) サンプル規模が大きくなるにつれて、母視聴率に近い値がサンプル視聴率として出てくる確率が高くなる

ただしサンプル集団が母集団と同じ大きさでも同じ数値になるとはかぎらない

←『復元抽出法』

4.3. 点推定

サンプル調査の精度では
pがわかっているものとして　サンプル視聴率x/nのどの値がどんな確率で出現するかを求めた

しかし実際のサンプル調査では
p は未知で
n と x が既知
知りたいのは母視聴率 p の方!

これを統計学では『推定』という
『推定』:「観察されたデータから確率変数の理論値を推定する」

視聴率調査での理論値は母視聴率(母比率)

サンプル視聴率 x/n からどうやって母視聴率 p を推測できるか?

やり方はいくつかあるが　よく使われるのは
『最尤推定法による点推定』
　『点推定』:特定の値の形で推定する
　『区間推定』:幅(区間)の形で推定する
では最尤推定とは?

−−具体的な例で説明しよう
今サンプル集団が 800 世帯、その視聴率が 40%、つまり 320 世帯がみていたとする
ここで知りたいのは関東地区全体の視聴率つまり母視聴率 p

⇒当てずっぽうではなく　ちゃんとした手がかりから推論するには?

既知のものから未知のものを知る上で、一番の手がかりは両者の関係　つまり観測値と理論値の間の関係
→これは二項分布ですでにわかっている
$f(x)={}_nC_x \cdot p^x \cdot (1-p)^{n-x}$

この場合は
$f(320)={}_{800}C_{320} \cdot p^{320} \cdot (1-p)^{480}$
ではこの式から p の値を推定するには?

いろいろ考え方があるが、一つのやり方は
<u>f(x)が確率を示す点に注目</u>
現実におきたことは「サンプル集団で 320 世帯がみていた」　とすればこれに最も近い状態になる p の値を p の推定値にすればいい
→「最もなりやすい=最も確率が高い」だから f(320)を最大にする p の値を求めればいい

直感的にいえば
<u>*現実に生起した事態は確率 1 だと考えられる⇒それに一番近いのは f(x)が最大になる状態*</u>

これを『最尤推定法』という
具体的な計算は　最大値は極値だから
$f(320)={}_{800}C_{320} \cdot p^{320} \cdot (1-p)^{480}$
を p の関数とみて　p で微分して=0 とおく

この関数を『尤度関数 likelihood function』という　ふつう L(p)と表記する
　確率関数 f(x)→p が定数で x が変数
　尤度関数 L(p)→x が定数で p が変数

計算結果だけ紹介すると　この場合は p=0.4 つまりサンプル視聴率と同じ値が母視聴率の最尤推定値　あたりまえに見えるが
重要なのは数学的な根拠があること!

最尤推定法の長所　(ゆるい条件つきだが)

1) 漸近不偏性:サンプル規模が大きくなるにつれて($n\to\infty$)、推定量 $\hat{a}$ の期待値が真の値になる　すなわち $E(\hat{a})=a$

2) 漸近有効性:サンプル規模が大きくなるにつれて、分散 $V(\hat{a})=E(\hat{a}-E(\hat{a}))^2$ が不偏推定値のなかで最小になる

3) 一致性:サンプル規模が大きくなるにつれて、推定値 $\hat{a}$ が真の値 a に一致する確率が 1 に近づく

短所

1) 上の長所が『頑健 robust』ではない:想定している確率関数(尤度関数)が少しちがうと成り立たない

● **推定値の良さの基準**

不偏性や有効性以外にも良さの基準はある

例えば『平均平方誤差 mean square error』:$E(\hat{a}-a)^2$　(真の値 a からの推定値 $\hat{a}$ のばらつき)が小さい

(なお有効推定値は<u>不偏推定値のなかで</u>平均平方誤差が最小　←分散 $E(\hat{a}-E(\hat{a}))^2$ が最小で不偏 $E(\hat{a})=a$ だから、$E(\hat{a}-a)^2$ が最小)

一つの推定量が多くの良さを兼ねそなえるとはかぎらない

例えば σ^2 の点推定では

1) 不偏推定値は $1/(n-1)\cdot\sum_i(x_i-m)^2$　(母平均と母分散がある全ての確率変数で)

2) 最尤推定値は $1/n\cdot\sum_i(x_i-m)^2$　(正規分布などの場合)

3) 平均平方誤差最小推定値は $1/(n+1)\cdot\sum_i(x_i-m)^2$　(正規分布の場合)

の 3 通りがある

漸近的には 1)2)3)は一致するが
サンプル規模が小さい場合にはどれかを選ぶことになる
それぞれに良し悪しが

1)は特定の分布を前提にしない点で頑健だが(←期待値の演算則だけで不偏性が証明できる)、くり返し測定をしない場合には不偏性自体が重要でない

3)は正規分布の(仮定をおく)場合には、実用性が高い (←直観的な「はずれにくさ」に一番近いのは平均平方誤差の小ささなので)

2)は数式で展開する上で良い性質をもつ
　例えば母分散の最尤推定値の平方根は母標準偏差の最尤推定値になる (母分散の不偏推定値の平方根は母標準偏差の不偏推定値にはならない!)
　←母数(=理論値)やデータの変換に関して不変 invariant

なお母比率の点推定でも
サンプル比率 x/n は母比率 p の最尤推定値で不偏推定値でもあるが
平均平方誤差が必ずしも最小にはならない

例えば p が 0.5 近くにある場合には
(x+1)/(n+2)の方が平均平方誤差が小さくなる
(=「(大きく)はずれにくい」)

4.4. 区間推定

点推定:特定の値すなわち点で推定
区間推定:「この幅(区間)に」の形で推定
大まかだが確からしさを数値的に評価できる

例えば

- **やり方その 1**

母視聴率 40%でサンプル集団 800 世帯なら
「サンプル集団では 38%から 42%の視聴率が出てくる」確率は 0.8
いいかえると
=「サンプル視聴率が母視聴率±2%の間にある」確率が 0.8
=「母視聴率-0.02≦サンプル視聴率≦母視聴率+0.02 である」確率が 0.8

「　」内に注目すると
「母視聴率-0.02≦サンプル視聴率≦母視聴率+0.02」は不等式の両辺を移項すると
「サンプル視聴率-0.02≦母視聴率≦サンプル視聴率+0.02」になる

したがって
=「サンプル視聴率-0.02≦母視聴率≦サンプル視聴率+0.02 である」確率が 0.8

サンプル視聴率 40%=0.4 を代入すれば
「0.4-0.02≦母視聴率≦0.4+0.02 である」確率が 0.8
=「38%≦母視聴率≦42%である」確率が 0.8
だといえる

ただし

上の区間推定にはずっと「母視聴率 0.4」という仮定が使われている　(←「Wald型」という)
この仮定を置かずに求めるには?

- **やり方その 2**

母比率 p の母集団から標本 1 個を抜き出して該当=1、非該当=0 とすると、これは
　E(･)=p
　V(･)=p(1-p)
の二項分布にしたがう
この確率変数からとった n 個のサンプルの値の平均、すなわち標本比率 x/n は中心極限定理によって
　E(･)=p
　V(･)=p(1-p)/n
の正規分布に漸近的にしたがう
(←ロンドンの出生比の検定と同じ考え方)

正規分布であることより　例えば
「x/n が E(･)±V(･)$^{0.5}$ 内」(=母平均から母標準偏差±1 つ分の間)にある確率は約 0.70←正規分布表で u=1.0 をみると 0.158 なので

あとはやり方その 1 と同じように　「x/n が E(･)±V(･)$^{0.5}$ 内」を　不等式にして変形していけばよいが　ここで二つのやり方にわかれる

やり方 2.1
　「x/n が E(･)±V(･)$^{0.5}$ 内」
=「x/n が p±{p(1-p)/n}$^{0.5}$ 内」
=「p-{p(1-p)/n}$^{0.5}$≦x/n≦p+{p(1-p)/n}$^{0.5}$」
=「x/n-{p(1-p)/n}$^{0.5}$≦p≦x/n+{p(1-p)/n}$^{0.5}$」
{p(1-p)/n}$^{0.5}$ は未知なので
{(x/n)(1-x/n)/n}$^{0.5}$ で代用して区間を求める

これは計算が楽だが　代用する際に p=x/n という形で p を実質的に推定している
→やり方その 1 と同じく Wald 型

それを避けるためには　中で p の値を代入せずに「　」の不等式を解けばよい!

で　やり方その 2.2
「x/n が $E(\cdot)\pm V(\cdot)^{0.5}$ 内」
→「$|x/n-E(\cdot)|\leqq V(\cdot)^{0.5}$」
→「$|x/n-p|\leqq(p(1-p)/n)^{0.5}$」
→「$(|x/n-p|)^2\leqq p(1-p)/n$」
→「$(x/n-p)^2\leqq p(1-p)/n$」
→「$(x/n)^2-2(x/n)p+p^2\leqq p(1-p)/n$」
→「$(x/n)^2-2(x/n)p+p^2-p(1-p)/n\leqq 0$」
→「$(x/n)^2-2(x/n)p+p^2-(1/n)p+(1/n)p^2\leqq 0$」
→「$(1+1/n)p^2-(2x/n+1/n)p+(x/n)^2\leqq 0$」

今　$(1+1/n)p^2-(2x/n+1/n)p+(x/n)^2=0$ の解を a, b とすると

「$(1+1/n)p^2-(2x/n+1/n)p+(x/n)^2\leqq 0$」
→「$a\leqq p\leqq b$」
と　p の範囲を示す不等式に書き換えられる!

だから、
「$a\leqq p\leqq b$」の確率は 0.70
といえる

a と b はサンプル規模 n とサンプル視聴世帯数 x で表されるから　p の区間を n と x だけで(=母視聴率の値を仮定せずに)推定できる

区間推定における
　推定された値のありうる範囲を『信頼区間』
　信頼区間にある確からしさを『信頼係数』
と呼ぶ

例えば
「サンプル視聴率 40%の場合､母視聴率の信頼係数 0.7 の信頼区間は 38~42%」(やり方 1)
「サンプル視聴率 40%の場合、母視聴率の信頼係数 0.7 の信頼区間は a~b」(やり方 2.2)

区間推定のやり方も複数ある　仮定の置き方がちがう　(ロンドンの出生比のときと同じ)
その 1:母視聴率の値を仮定
その 2.1:正規分布と母視聴率の値を仮定
その 2.2:正規分布を仮定
(←中心極限定理からは*漸近的に*正規分布にしたがうといえるだけ)

統計的推定とは　一定の仮定をおくことで数量的にその良さが評価できる形で推定する
それによって
　① どんな仮定を置いたかが明確になる
　② 仮定の負荷を減らす方向に改良できる
(→ことによってより良い推定ができる)

例えば
やり方その 1 や 2.1 などの Wald 型は計算が楽だが　推定すべき母視聴率を x/n で代入する
→サンプル規模が小さい場合や母比率が 0 や 1 に近い場合には精度が悪くなる
やり方その 2.2 は計算の手間が少しかかるが　推定すべき母視聴率を仮定しないでいい
→その 1 や 2.1 ほど精度が落ちない

- **信頼係数の意味**

区間推定は実用性が高い
例えばデータが得られた後で、統計的な誤差を考慮して判断するときは、検定ではなく、区間推定が第一選択とされる

けれども信頼係数の意味が大きな問題になってきた

信頼係数 0.8 とはどういうこと?
本来の(=現在の標準的な)定義では
「この範囲にあると考えて、10 回中 8 回は正しい」(=1 回の推定では正しいとも正しくないともいえない)
←母数(例えば母視聴率)は(未知だけれども)決まった数値だと定義されているので
推定した区間の内にある(=正しい)か内にない(=正しくない)かのどちらか

ところが実際には　信頼係数は
「この区間内にある確からしさが 0.8」(=1 回の推定の正しさの程度)
と解釈されている

定義と実際の解釈がずれてきた(=統計学的には「嘘」の解釈で使われてきた)
→
現在ではベイズ統計学の考え方を導入するようになりつつある(=信頼係数はベイズ統計学的な意味での確率だとする)

ベイズ統計学:事前分布と事後分布という形で母数も確率変数としてあつかう
この講義では直接はふれないが、標準的な考え方とベイズ統計学の考え方とが重なりやすい形で解説している

4.5. t 分布をつかった区間推定の例

仮定の負荷を減らせる場合は他にもある
例えば
観測値の出方が正規分布にしたがう、つまり正規分布にしたがう確率変数の観測値から母平均 μ を推定する場合
t 分布をつかって推定するのが一般的

この推定法の長所は
1) 他の仮定がいらない
2) 頑健性がある　(正規分布からずれていても妥当性がある程度ある)

詳しくいえば−−
正規分布にしたがう確率変数の観測値が n 個ある場合、母平均を μ、標本平均を m、標本分散を s^2 とすると、

$(m-\mu)/\{s/(n-1)^{0.5}\}$が『自由度 n-1 の t 分布』にしたがう
この$(m-\mu)/\{s/(n-1)^{0.5}\}$を『t 統計量』とよぶ(t 統計量は t 分布にしたがう確率変数)

自由度 n-1 の t 分布がどの幅の値をどんな確率でとるかは t 分布表からわかる
例えば
n が 10 個ならば自由度 9 の t 分布
t 分布表を見ると自由度 $v=9$ の行で $\alpha=0.15$ の値は 1.1
→$(m-\mu)/\{s/(n-1)^{0.5}\}$が+1.1 より大きくなる確率は 0.15

したがって
「$-1.1 \leqq (m-\mu)/\{s/(n-1)^{0.5}\} \leqq +1.1$」になる確率は $1-0.15\times2=0.7$ になる

あとは「　」のなかを変形していけばよい

「$-1.1 \leqq (m-\mu)/\{s/(n-1)^{0.5}\} \leqq +1.1$」
→「$-1.1\{s/(n-1)^{0.5}\} \leqq (m-\mu) \leqq +1.1\{s/(n-1)^{0.5}\}$」
→「$-m-1.1\{s/(n-1)^{0.5}\} \leqq -\mu \leqq -m+1.1\{s/(n-1)^{0.5}\}$」
→「$m-1.1\{s/(n-1)^{0.5}\} \leqq \mu \leqq m+1.1\{s/(n-1)^{0.5}\}$」

つまり
母平均 μ の、信頼係数 0.7 の信頼区間は
$m-1.1\{s/(n-1)^{0.5}\} \sim m+1.1\{s/(n-1)^{0.5}\}$
である

数値例でいうと　例えば
観測値が 10 個で
1.4, 1.8, 0.9, 2.1, 1.7, 1.8, 1.5, 1.7, 1.8, 1.9
だとすると

$m=1.66$
$s^2=0.0984$　したがって $s=0.3137$
$1.1\{s/(n-1)^{0.5}\}=1.1 \times 0.3137/(9^{0.5})=0.115$

つまり
母平均 μ の信頼係数 0.7 の信頼区間は
1.66 ± 0.115 すなわち $1.545 \sim 1.775$

で区間推定の話は終わりだが
もう一度不等式に戻ってみよう

「$-1.1 \leqq (m-\mu)/\{s/(n-1)^{0.5}\} \leqq +1.1$」になる確率は 0.7　という命題は
「$\mu-1.1\{s/(n-1)^{0.5}\} \leqq m \leqq \mu+1.1\{s/(n-1)^{0.5}\}$」になる確率が 0.7　という命題にも変換できる

区間推定では　観測値 m を既知として理論値 μ を考えた
それに対して、理論値 μ について仮説があれば、<u>その仮説が現実の観測値 m にあてはまるかどうかを上の不等式で評価できる</u>

先の数値例をそのまま使うと
観測値が 10 個で 1.4, 1.8, 0.9, 2.1, 1.7, 1.8, 1.5, 1.7, 1.8, 1.9 だとする

母平均 $\mu=1.5$ という仮説(H_0)の下でこのデータ(D)は出やすいといえるか?
(←Arbuthnot の判定と同じ考え方)

データから s と n は計算できる
$s^2=0.0984$　だから $s=0.3137$
$1.1\{s/(n-1)^{0.5}\}=1.1\times0.3137/(9^{0.5})=0.115$
したがって
「$\mu-0.115 \leqq m \leqq \mu+0.115$」になる確率が 0.7
それゆえ　$\mu=1.5$ が正しければ
「$1.485 \leqq m \leqq 1.615$」になる確率が 0.7

観測値の平均 m は 1.66 だった　つまりこのデータは $\mu=1.5$ という仮説の下では出にくい
したがって
$\mu=1.5$ という仮説はデータにあてはまらない
と判定できる
⇒<u>統計的検定の考え方は区間推定と裏表</u>

5. 仮説の検定
5.1. 基本的な考え方

推定:データから母集団の値を推定
検定:母集団に関する予想をデータで検証

原型はやはり Arbuthnot の判定
H_0:男性が多い年と女性が多い年は 1/2 の確率で生起する
D:男性が多い年が 82 年続いている

H_0 の下で D が出現する確率は$(1/2)^{82}$
この値はきわめて小さいから　仮説 H_0 は正しいとはいえないと判断する
⇒偶然の誤差を考えた上での判定

考え方はこのまま　あとは方法を洗練する(より良い性質をもつ方法にする)だけ!

具体的にどうするかというと

H_0:母視聴率が 0.4
D:サンプル規模 200 でサンプル視聴率 0.37

このとき仮説 H_0 は外れと判断できるか?

考え方:一定の基準でみて　H_0 の下では D は出にくいならば　「母視聴率が 0.4」という仮説は×であると判定する

では「サンプル視聴率 0.37 は出にくいか?」

これにもやり方は複数ある!

- **やり方その 1**

D「サンプル規模 200 でサンプル視聴率 0.37」が出る確率を直接計算する
=0.04

この形だと出にくさは評価しにくい
例えば
H_0「0.41」なら D の出る確率は 0.03
H_0「0.39」なら D の出る確率は 0.05

……判断に困る
別のやり方は?

- **やり方その 2**

区間推定のように考えて
1) *事前に*、「母視聴率が 0.4」という仮説が正しければ、サンプル視聴率の値として出てもおかしくない区間を置いて
2) *実際に調べた*サンプル視聴率がその区間の外ならば「仮説は×」　区間の内なら「仮説は×でない」と判断する

例えば
区間として「36~44%」を置けば
もしサンプル視聴率が 46%なら区間外なので、仮説は×
もしサンプル視聴率が 43%なら区間内なので、仮説は×でない
とする　ここで注意すべき点!
「×か×でないか」を判定するもので、「○か○でないか」ではない

区間をどう置けばいいか?
素朴に考えて　区間が狭ければ「×」と判定しやすい

例えばもし「母視聴率が 0.4」という仮説が正しければ　サンプル規模 200 だと

区間 A「36~44%」なら　区間内にサンプル視聴率が入る確率は 0.75
⇒実際のサンプル視聴率が区間内でなかったら仮説は×だと判定して、それがまちがいである確率は 0.25

区間 B「34~46%」なら　区間内にサンプル視聴率が入る確率は 0.90
⇒実際のサンプル視聴率が区間内でなかったら仮説は×だと判定して、それがまちがいである確率は 0.1

A か B かはその判定にどれだけのリスクを許すかで決めればよい(決めるしかない)
例えば新薬の効果を調べるなら　H_0「差はない」が本当は○なのに　「×である」とする確率はかなり小さい方がよいだろう(副作用や経済的コストを考えると)

仮説が本当は○なのに「×である」と判定する確率:『有意水準』あるいは『危険率』
例えば　区間 A は危険率 0.25　区間 B は危険率 0.1　になる

以上まとめると
1) 仮説を立てる
2) *事前に*、一定の危険率を見込んで区間を決める
3) *実際に調べた*値が 2)の区間に入らなければ「仮説は×である」と判定する

5.2. 有意水準と検出力

5.1 で紹介した形ではまだ不十分

→

危険率 0.25(仮説が正しければその区間内の値をとる確率 0.75)の区間は複数ある!

例えば

「サンプル視聴率が 36~44%」

「サンプル視聴率が 0~42%」

どちらの区間にも確率 0.75 でおちる

☆区間推定の信頼区間も本当は複数ある!

危険率だけならどちらも同じ

でも直感的には　「サンプル視聴率が 36~44%」の方が「サンプル視聴率が 0~42%」よりも良い幅のとり方だと思える

なぜ良いと思えるのだろうか?

サンプル規模 200 で「母視聴率 0.4」という仮説を危険率 0.25 で検定するとしよう

区間 A の置き方は複数ある　例えば

区間 A1:サンプル視聴率が「36%」~「44%」

区間 A2:サンプル視聴率が「0%」~「42%」

区間 A3:サンプル視聴率が「37%」~「100%」

では三つの区間の置き方は何がちがう?

→

「母視聴率 0.4」という仮説が本当はまちがっている場合

にちがいが出てくる

例えば

本当は母視聴率 0.34 だったりすると

区間 A1 に入る確率は 0.25
⇒「母視聴率 0.4」という仮説は本当は×なのに「×でない」とする確率は 0.25
区間 A2 に入る確率は 0.99
⇒「母視聴率 0.4」という仮説は本当は×なのに「×でない」とする確率は 0.99
区間 A3 に入る確率は 0.17
⇒「母視聴率 0.4」という仮説は本当は×なのに「×でない」とする確率は 0.17

→区間 A3 が一番まちがう確率が低い　それゆえ best な置き方

母視聴率 0.46 だったりすると

区間 A1 に入る確率は 0.31
⇒「母視聴率 0.4」という仮説は本当は×なのに「×でない」とする確率は 0.31
区間 A2 に入る確率は 0.14
⇒「母視聴率 0.4」という仮説は本当は×なのに「×でない」とする確率は 0.14
区間 A3 に入る確率は 0.99
⇒「母視聴率 0.4」という仮説は本当は×なのに「×でない」とする確率は 0.99

→区間 A2 が一番まちがう確率が低い　それゆえ best な置き方

区間をどこに置くかで
<u>仮説が本当は×だった場合　×なのに「×でない」とする確率</u>
がちがってくる
この「仮説が本当は×なのに「×でない」とする」ことを「第二種の誤り」という

『第一種の誤り』:仮説は本当は○なのに「×である」とする誤り
『第二種の誤り』:仮説が本当は×なのに「×でない」とする誤り
危険率は第一種の誤りの確率にあたる　これをふつう α で表す
第二種の誤りの確率はふつう β で表す　特に $1-\beta$ を『検出力 power』と呼ぶ

5.3. 最強力検定の考え方

検定では事前に(つまり調べる前に)区間の大きさと置き方をきめる
⇒第一種と第二種両方の誤りを考慮する必要

一番良い区間とは?
「第一種の誤りの確率(=危険率)が一定の値 α 以下で、第二種の誤りの確率 β が最小(=検出力 1-β が最大)の区間」と考えればよい

検出力の大きさは母視聴率の値でかわる
*母視聴率がある値の下で*検出力が最大
=『最強力 most powerful』

例えば母視聴率 0.34 なら区間 A3 が、母視聴率 0.46 なら区間 A2 が最強力
検出力の大きさは母視聴率の値で変わる

⇒検出力は母視聴率を定義域とする関数　になっている
これを『検出力関数』という

*母視聴率がどの値でも*検出力が最大
=『一様最強力 uniformaly most powerful』

一様最強力な検定はつねにあるわけではない
(例えば A1、A2、A3 の場合にもない)
ない場合によく使うのが『一様最強力不偏 uniformaly most powerful unbiased』検定　すなわち　『不偏検定』:検出力 1-β が危険率 α を下回らない検定法　のなかで一様最強力
先の例では A1 が一様最強力不偏検定

これがどんな意味で良い検定かというと
① A2 や A3 は母視聴率の値によっては A1 より強力だが、検出力が 0.25 を下回ることも
② (①の言い換えだが)A1 の検出力は仮説の値 0.4 の近くで低く、遠くで高い
→
本当の母視聴率が仮説の 0.4 に近い場合は仮説を「×ではない」と判定しやすく遠い場合には「×ではない」と判定しにくい
いわば「<u>当たらずとも遠からず</u>」!

A2 や A3 ではそうはかぎらない
だから本当の母視聴率が見当つかない場合には区間 A1 型が良い

まとめると　統計的検定とは

1) 仮説をたてる
2) *<u>事前に</u>*、危険率 α を設定して区間の大きさを、検出力 1-β を考慮して区間の位置を決める
　A1 型の置き方を『両側検定』とよぶ
　A2 や A3 型の置き方を『片側検定』とよぶ
実際には特に理由のない限り『両側検定』、仮説の値以上か以下かどちらかにしかならないなどの特定の場合に『片側検定』
3) *<u>実際に調べた</u>*データでの値が 2)での区間の外なら「仮説は×」と判定する

5.4. いろいろな検定法

検定とは何か　をおさらいしておくと

<u>**ある仮説**</u>　が正しければ
<u>**観測値から求められる量**</u>　が
<u>**特定の確率分布**</u>　にしたがう確率変数になる

ことを使って　その量の出現確率の大きさで仮説の正しさを評価する手続き
この「観測値から求められる量」を検定統計量とよぶ

例えば　ロンドンの出生比の話でいえば
<u>**ある仮説**</u>:*「出生児が男性の確率 p が 0.5」あるいは「p が 0.52」*
が正しければ
<u>**観測値から求められる量**</u>:$1/N\sum z_i$
が
<u>**特定の確率分布**</u>:$E(\cdot)=np$、$V(\cdot)=\{np(1-p)\}/N$ *の正規分布*
にしたがう確率変数になる
ことを使って　その量 $1/N\sum z_i$ の出現確率の大きさで仮説「p が 0.5」や「p が 0.52」の正しさを評価する手続き

母視聴率の検定の例では
仮説:母視聴率つまり母集団での比率
検定統計量:サンプル視聴率
で
特定の確率分布:サンプル視聴率が二項分布、または近似的に正規分布にしたがう
ことを使って
その出現確率の大きさで正しさを評価した

仮説の種類に応じてどの量を使うか(=何が検定統計量になるか)は変わってくる
だから実際に使う上では

<u>仮説の種類と検定統計量(+それがどんな場合にどんな確率分布にしたがうか)の組み合わせ</u>

を知っておく必要がある(対応する検定統計量がない場合には仮説は検定できない)

例えば––

(1) 母比率の検定

これは説明済み
仮説の種類は「母集団のある変数の比率 p」
検定統計量は「サンプル集団でのその変数の比率 x/n」
<u>サンプル集団が無作為抽出でつくられた場合</u>、この検定統計量は二項分布　もしくは中心極限定理より近似的に正規分布にしたがう

(2) 母平均の検定

仮説の種類は「母集団のある変数がとる平均値 μ」
検定統計量は「$(m-\mu)/\{s/(n-1)^{0.5}\}$」(仮説により μ の値があたえられるので後はデータだけから計算できる)
<u>n 個の観測値が正規分布にしたがって出現する場合</u>、この検定統計量は自由度 n-1 の t 分布にしたがう

詳しくいうと––
母平均を μ、標本平均を m、標本分散を s^2 で表すと、
$(m-\mu)/\{s/(n-1)^{0.5}\}$が自由度 n-1 の t 分布にしたがう

たんに「t 検定」というとこれを指すことが多い

- **実際の使い方**

4.5 の例をそのまま使って

観測値が 1.4, 1.8, 0.9, 2.1, 1.7, 1.8, 1.5, 1.7, 1.8, 1.9 のデータで 「母平均 μ=1.4」という仮説を危険率 0.1 で検定してみると

観測値が 10 個　なので　$(m-\mu)/\{s/(n-1)^{0.5}\}$は自由度 9 の t 分布にしたがう

特に理由がないので両側検定(区間 A1 型)

→t 分布表での自由度 9 の危険率 0.1 の値をみると 1.833　(ν=9、2α=0.10 のところ)

したがって棄却域(そこに入ると仮説は×だと判断する区間)は 1.833 以上と -1.833 以下

あとは計算するだけ

$m=1.66$、$s^2=0.0984$　したがって $s=0.3137$

$s/(n-1)^{0.5}=0.3137/(9^{0.5})=0.1046$

$\mu=1.4$ という仮説の下では

$(m-\mu)/\{s/(n-1)^{0.5}\}=(1.66-1.4)/0.1046=2.49$

で棄却域に入る

→$\mu=1.4$ という仮説は棄却される

もしこれが $\mu=1.6$ だという仮説であれば

$(m-\mu)/\{s/(n-1)^{0.5}\}=(1.66-1.6)/0.1046=0.57$

だから棄却域に入らない

→$\mu=1.6$ という仮説は棄却されない

(3) 母分布に関する検定

仮説の種類は「母集団のある変数の値を k 個のカテゴリーに分類した場合の各カテゴリーの比率の分布 $p_1, p_2, \ldots, p_k$ ($\sum_i p_i=1$)」
検定統計量は「$\sum_i\{(f_i-np_i)^2/np_i\}$」 ($f_i$ は、n 個の観測値を上の各カテゴリーに分類した場合にそれぞれに該当する個数で、$\sum_i f_i=n$)
←仮説から p_i があたえられるので後はデータから計算できる

個数の分布 $f_1, f_2, \ldots, f_k$ が母平均 np_i、母分散 $np_i(1-np_i)$、母共分散 $-np_ip_j$ の多次元正規分布にしたがって出現する場合、この検定統計量は『自由度 k-1 のカイ 2 乗分布』にしたがう
−−まとめていえば、
母分布が $p_1, p_2, \ldots, p_k$ ($\sum_i p_i=1$)、観測値の個数が $f_1, f_2, \ldots, f_k$ ($\sum_i f_i=n$)のとき、
$\sum_i\{(f_i-np_i)^2/np_i\}$は自由度 k-1 のカイ 2 乗分布にしたがう
(この $\sum_i\{(f_i-np_i)^2/np_i\}$は『(適合度の)カイ 2 乗統計量』とよばれる)

「$f_1, f_2, \ldots, f_k$ が~の多次元正規分布にしたがう」にあてはまる具体例としては例えば
サンプル集団が無作為抽出で(=サンプルになる確率が等しいよう)つくられた場合、
母集団でのある変数の比率の分布を $p_1, p_2, \ldots, p_k$ ($\sum_i p_i=1$)、サンプル集団での対応する個数の分布を $f_1, f_2, \ldots, f_k$ ($\sum_i f_i=n$)とすると、
$\sum_i\{(f_i-np_i)^2/np_i\}$は漸近的にカイ 2 乗分布にしたがう (←中心極限定理により、サンプル比率は漸近的に正規分布にしたがうように)

①
2 種類の分布が出てくることに注意
1) 観測値の値にあたるものの母集団での比率の分布 $p_i(=p_1, p_2, \ldots, p_k)$と　サンプル集団での個数の分布 $f_i(=f_1, f_2, \ldots, f_k)$
2) サンプル集団での個数の分布 f_i の出現の仕方がしたがう確率分布

つまり　適合度検定では、観測された値(を k 種類に分類した場合)の個数の分布 f_i を確率変数として考えており、p_i はその確率変数がしたがう確率分布の理論値にあたる

②
サンプル集団が無作為抽出でつくられた場合、厳密には、個数の分布 f_i は p_i を理論値とする<u>多項分布(=二項分布を多変数に拡張した分布)にしたがう</u>

多項分布の確率関数は
$f(x_1, x_2, \ldots, x_k)$
$=\{n!/(x_1!x_2!\ldots x_k!)\}(p_1)^{x1}(p_2)^{x2}\ldots(p_k)^{xk}$
なお x!は x の階乗(ただし 0!=1)
例えば　袋のなかに赤球が 0.2、青球が 0.3、黒球が 0.3、白球が 0.2 の比率で入っている場合、6 個球を取り出して、赤 1、青 2、黒 2、白 1 になる確率は
$f(1, 2, 2, 1)$
$=\{6!/(1!2!2!1!)\}(0.2)^1(0.3)^2(0.3)^2(0.2)^1$

<u>n が十分大きければ、多項分布は多次元正規分布(=正規分布を多変数に拡張した分布)に一致する</u>　という関係にある
→
$\sum_i\{(f_i-np_i)^2/np_i\}$は<u>漸近的に</u>カイ 2 乗分布にしたがう

- **実際の使い方その 1:適合度の検定**

メンデルの仮説(メンデルの法則)
「豆の黄/緑と丸い/皺皺はそれぞれ 3:1 の比率で出現して相互が独立である」
具体的にいえば「{黄·丸、黄·皺、緑·丸、緑·皺}の母分布は{9/16, 3/16, 3/16, 1/16}」
実際にとれた豆は{黄·丸、黄·皺、緑·丸、緑·皺}={315, 101, 108, 32}
とれた豆は無作為に出現したものとして　メンデルの仮説を危険率 0.1 で検定してみると
観測値の値の分布の個数 k は 4 なので
$\sum_i (f_i-np_i)^2/np_i$ は自由度 3 のカイ 2 乗分布にしたがう
また　$\sum_i (f_i-np_i)^2/np_i$ は正の値しかとらないので片側検定
→カイ 2 乗分布表で自由度 3 の危険率 0.1($\nu=3, \alpha=0.10$)の値をみると 6.25
したがって　棄却域は 6.25~
あとは計算するだけ
母分布 p_i は{9/16, 3/16, 3/16, 1/16}、観察された分布 f_i は{315, 101, 108, 32}だから
$\sum_i (f_i-np_i)^2/np_i=0.470$
だから棄却域には入らない
→メンデルの仮説{9/16, 3/16, 3/16, 1/16}は棄却されない
実は危険率 0.9 でも(つまり本当は○なのに×とする可能性が 90%という、きわめて棄却しやすい基準でも)メンデルの仮説は棄却されない
データにとてもよくあてはまる仮説だった
……教科書的にはそういうことだが　このケースはなにか違和感がある
理由その 1:出来過ぎのデータ?
かえって怪しげ　→捏造の疑いアリ　真実は神のみぞ知るだが…
理由その 2:仮説検定の論理との食い違い?
検定とは本来、ある仮説を正しいと考えた場合、「大目にみても」あり得ない結果がデータで出ていることを示すことで否定する手続き
　Arbuthnot の判定もそう
⇔メンデルの仮説のケースは肯定したい仮説
肯定したい仮説のあつかい方には現状ではいろんな立場があり　統一見解はないが
検出力関数で考えると　ある程度根拠づけられる　(→5.5)

- **実際の使い方その 2:独立性の検定**

サンプル集団で変数 $X(x_i)$と変数 $Y(y_j)$のクロス集計の値が f_{ij} だった場合、母集団において変数 X と変数 Y が独立かどうかを検定できる

「独立である」という仮説が正しければ、母集団における値の分布 p_{ij} について

$p_{ij}=(f_{i\cdot}/n)\times(f_{\cdot j}/n)=(f_{i\cdot}\times f_{\cdot j})/n^2$

が成り立つ(後述)

それゆえ、

$\sum_{ij}(f_{ij}-np_{ij})^2/np_{ij}$

$=\sum_{ij}(f_{ij}-f_{i\cdot}f_{\cdot j}/n)^2/(f_{i\cdot}f_{\cdot j}/n)$

はカイ 2 乗分布にしたがう

ただし、この時の自由度は ij-1 ではなく*(i-1)(j-1)*になる

例えば

男性か女性かと死後の世界を信じるかどうかは関係あるかを調べたところ、サンプル調査では次のような結果になった

	信じる	信じない
男性	350	100
女性	400	150

このとき「性別と死後の世界を信じる信じないは独立である」という仮説を危険率(有意水準)0.05 で検定してみよう

性別を変数 X(1=男性, 2=女性)、信じるかどうかを変数 Y(1=信じる, 2=信じない)とすると

上の表は

$f_{11}=350,\ f_{12}=100,\ f_{21}=400,\ f_{22}=150$

という値の分布になっている

このとき

$f_{1\cdot}=f_{11}+f_{12}=450,\ f_{2\cdot}=f_{21}+f_{22}=550$

$f_{\cdot 1}=f_{11}+f_{21}=750,\ f_{\cdot 2}=f_{12}+f_{22}=250$

だから

$\sum_{ij}(f_{ij}-np_{ij})^2/np_{ij}$

$=\sum_{ij}\{(f_{ij}-f_{i\cdot}f_{\cdot j}/n)^2/(f_{i\cdot}f_{\cdot j}/n)\}$

$= (f_{11}-f_{1\cdot}f_{\cdot 1}/n)^2/(f_{1\cdot}f_{\cdot 1}/n)+(f_{12}-f_{1\cdot}f_{\cdot 2}/n)^2/(f_{1\cdot}f_{\cdot 2}/n)$

$+(f_{21}-f_{2\cdot}f_{\cdot 1}/n)^2/(f_{2\cdot}f_{\cdot 1}/n)+(f_{22}-f_{2\cdot}f_{\cdot 2}/n)^2/(f_{2\cdot}f_{\cdot 2}/n)$

$=3.37$

自由度 1 のカイ 2 乗分布の危険率(有意水準)0.05 の値は 3.84、

→独立であるという仮説は棄却されない　すなわち独立でないとはいえない

- **独立性の検定と適合度の検定のちがい**

独立性の検定は適合度の検定の特殊ケース

どこがちがうか?

通常の適合度の検定は理論値が既知

　例えば　メンデルの仮説が正しければ色と形の比率はこうなる　など

独立性の検定では理論値は未知

　独立という仮説の下で最尤推定法によって母集団の値の分布を求めると

$p_{ij}=(f_{i\cdot}/n)\times(f_{\cdot j}/n)$

が成り立つ

ただし　この推定の際に p_{ij} の値のとり方にある条件をあたえる

そのため自由度が<u>*(i-1)(j-1)*</u>になる

- **実際の使い方その 3:ポアソン分布かどうか**

二項分布:ある事象が起きる確率が p の場合に、n 回やってみてその事象が x 回起きる確率

n が大きい場合は計算が大変　→正規分布で近似するが、p がとても小さい場合はあまりよい近似にならない
→こうした場合は『ポアソン分布』を使う

ポアソン分布とは　確率関数が
$f(x)=(e^{-\lambda}\lambda^{x})/x!$
λ は理論値(母数) parameter
e は自然対数の底 2.7182...
x!は x の階乗(ただし 0!=1)
になる離散型の確率分布で、
$E(X)=V(X)=\lambda$
である

ある観測単位(例えば特定の観測時間)内の試行回数が n、各試行でその事象が起きる確率を p とすると、その観測単位でその事象が x 回起きる確率は、近似的に
$f(x)=(e^{-np}(np)^{x})/x!$
すなわち、$\lambda=np$ のポアソン分布にしたがう

この近似の利点
(a) 計算が楽
(b) np を一つの値として扱える
例えば、np の最尤推定値は x の平均値、すなわち各観測単位で起きた回数の平均になる(n もしくは p が具体的にわからなくても np ならわかることがある)

これもかなり役立つ知識
例えば

「ごく稀におきるが(重大な意味がある)事件や事故が本当にただの偶然なのか、それとも何らかの恒常的なしくみの下で起きているのか」
を識別できる

歴史上最初の例
「プロイセンの騎兵連隊における馬に蹴られて死亡した兵士の数」
L.Bortkewicz による

1870 年までの 20 年間に、プロイセンの 10 個の騎兵連隊で、1 年間に馬に蹴られて亡くなった兵士の数を調べると

死亡数	0	1	2	3	4	5
観測数	109	65	22	3	1	0

1 連隊での 1 年間の平均死亡数 0.61

これを np の推定値として $\lambda=0.61$ のポアソン分布にあてはめると

死亡数	0	1	2	3	4	5
観測数	109	65	22	3	1	0
あてはめ	108.6	66.3	20.2	4.1	0.6	0.1

観測数にきわめてよくあてはまっている!
どのくらいあてはまっているか?

→カイ 2 乗分布による適合度の検定
「あてはめ」を理論分布とすると
カイ 2 乗値$\sum_i(f_i-np_i)^2/np_i=0.78$
自由度 5 のカイ 2 乗分布は $\alpha=0.975$ で 0.83

→危険率 0.975 というきわめてきびしい基準でみても、「$\lambda=0.61$ のポアソン分布にしたがう」という仮説は×ではない

$\lambda=np$ が 0.61 で、20 年間一定だとすれば、
もしこの 20 年間に、騎兵連隊の規模や馬の飼育法に大きく変化がなかったとすれば、
n は(具体的な数値はわからないが)一定だと考えられる
→p は(具体的な数値はわからないが)一定だと考えられる
→馬に蹴られる死亡事故はある規則性をもって発生している
→何か恒常的なしくみがある　と考えられる

- **カイ 2 乗検定がよく使われる理由**

何らかのしくみ
↓
理論値の分布 $\{p_1, p_2, \ldots, p_k\}$
↓無作為に出現する場合
データの個数の分布 $\{f_1, f_2, \ldots, f_k\}$
　→適合度のカイ 2 乗検定は$\boxed{\text{ここ}}$を検証

⇒理論値(母数)が分布の形をとる仮説が検証できる　(理論値がポアソン分布にしたがう、とか　多項分布の理論値はこれこれ　とか)
←統計学はもともと分布をあつかう

(4) 母分散に関する検定

仮説の種類は「母集団のある変数がとる分散 σ^2 の値」
検定統計量は「$(ns^2)/\sigma^2$」(仮説から σ^2 があたえられるので後はデータから計算できる)
n 個の観測値がそれぞれ正規分布にしたがって出現する場合、この検定統計量は『自由度 n-1 のカイ 2 乗分布』にしたがう

くり返しになるが––
母分散を σ^2、標本の分散を s^2 とすると
$(ns^2)/\sigma^2$ は自由度 n-1 のカイ 2 乗分布にしたがう

これも『カイ 2 乗検定』の一つ

- **実際の使い方その 1**

4.5 で使った数値例をそのまま使えば
観測値が 1.4, 1.8, 0.9, 2.1, 1.7, 1.8, 1.5, 1.7, 1.8, 1.9 の 10 個だとする
このデータの下で「母分散 $\sigma^2=0.09$」という仮説を危険率(有意水準)0.1 で検定してみよう

観測値が 10 個　なので　$(ns^2)/\sigma^2$ は自由度 9 のカイ 2 乗分布にしたがう
特に理由がないので両側検定(区間 A1 型)

→カイ 2 乗分布表をみるわけだが
カイ 2 乗分布は左右対称ではないので、両側検定の場合には　自由度 9 で危険率 0.95(v=9, α=0.95)の値と、自由度 9 で危険率 0.05(v=9, α=0.05)の値をみる

実際に分布表をみると　$(v=9, \alpha=0.95)$の値は 3.325、$(v=9, \alpha=0.05)$の値は 16.919 だから
棄却域(そこに入ると仮説は×だと判断する区間)は~3.325 と 16.919~

あとは計算するだけ
$n=10$、$s^2=0.0984$　したがって $ns^2=0.984$
「$\sigma^2=0.09$」という仮説の下では
$(ns^2)/\sigma^2=0.984/0.09=10.93333$
で棄却域に入らない
→$\sigma^2=0.09$ という仮説は棄却されない

- **実際の使い方その 2**

この検定統計量は区間推定にも使える

同じ数値例を使うと
観測値が 1.4, 1.8, 0.9, 2.1, 1.7, 1.8, 1.5, 1.7, 1.8, 1.9 というデータの下で、母分散 σ^2 を信頼係数 0.9 で区間推定するとすれば

$(v=9, \alpha=0.95)$の値は 3.325、$(v=9, \alpha=0.05)$の値は 16.919 だから

$3.325<(ns^2)/\sigma^2<16.919$
$1/16.919<\sigma^2/(ns^2)<1/3.325$
$ns^2/16.919<\sigma^2<ns^2/3.325$
$ns^2=0.984$ より
$0.984/16.919<\sigma^2<0.984/3.325$
$0.0582<\sigma^2<0.296$

→このデータの下では、母分散 σ^2 は信頼係数 0.9 で 0.0582~0.296 の区間にあるといえる

母分散の検定や区間推定は実際にはあまりやらないが、こちらの方がカイ 2 乗分布を使った検定の基本形

5.5. 検出力関数の意味

検出力関数とは何か?
⇒この関数は検出力とともに危険率を表す

…意外に思うかもしれないが

「本当の母視聴率が 0.4001 である場合のサンプル視聴率の出方」≒「本当の母視聴率が 0.4 である場合のサンプル視聴率の出方」
である以上

「本当の母視聴率が 0.4001 である場合に、『仮説を×と判定する』区間にサンプル視聴率が入る確率」≒「本当の母視聴率が 0.4 である場合に、『仮説を×と判定する』区間にサンプル視聴率が入る確率」
になるはず

このとき、≒の左側の「本当の母視聴率が 0.4001 である場合に、『仮説を×と判定する』区間にサンプル視聴率が入る確率」とは
⇒
本当の母視聴率が 0.4001 だから、仮説は×である場合に「仮説は×だと判定する」確率
すなわち、母視聴率が 0.4 という仮説の下で、本当の母視聴率が 0.4001 の場合に第二種の過誤をしない確率=**検出力**　にあたる

≒の右側の「本当の母視聴率が 0.4 である場合に、『仮説を×と判定する』区間にサンプル視聴率が入る確率」とは
⇒
本当の母視聴率が 0.4 だから、仮説は×でない場合に「仮説は×だと判定する」確率
つまり、母視聴率が 0.4 という仮説の下で、第一種の過誤をする確率=**危険率**にあたる

したがって
本当の母視聴率 θ が θ_0 にごく近い場合には
<u>検出力 1-β≒危険率 α</u>になる

それゆえ　検出力関数の値 B(θ)は
<u>検出力と危険率を同時に表す</u>

変な感じがするかもしれないが

母視聴率が 0.4 という仮説の下では、本当の母視聴率がどこにあろうと、「仮説は×」だと判定する区間の置き方は同じ
だから
「本当の母視聴率が仮説の値にごく近い場合に、その区間にサンプル視聴率が入る確率」と「本当の母視聴率が仮説の値である場合に、その区間にサンプル視聴率が入る確率」はほとんど同じになる

検出力関数のありがたみ　はどこかというと

その形状をみれば　危険率と検出力がどんな関係にあるかがわかる
だから、検定結果からどんな判断ができるかも見通しがつく

具体的にいうと
一般的にいって、良い検定方法というのは
1）危険率 α はできるだけ低い(例えば 0.05 とか 0.01 とか)
2）検出力 1-β はできるだけ高い(つまり 1.0 に近い)
という二つの性質をみたすもの

ただし　これまで述べたように　本当の母視聴率 θ が仮説の値 θ_0 にごく近い場合には　<u>検出力 1-β≒危険率 α</u>　になる

すなわち<u>**検出力関数 B(θ)は(θ_0, α)を必ず通る**</u>
したがって、仮説の値 θ_0 のごく近くでは検出力 1-β は必ず α まで落ち込んでしまう
⇒危険率を低くすると検出力も低くなる!
それゆえ 1)と 2)をみたす良い検定方法とは
<u>**検出力関数 B(θ)が仮説の値 θ_0 の近くで急激に凹む**</u>形のもの
凹み方は急激であればあるほどよい

凹み方をきめる要因は三つ
a) 仮説が本当は×の場合に統計量がしたがう確率分布=「非心分布」の種類
b) 危険率 α
c) サンプル規模 n

さらに
現実のデータに検定を使う場合は、凹み方が急激であることにさらに大きな意味が加わる
というのは
もし θ_0 の近くで B(θ)が急激に凹であれば
<u>仮説検定で「×でない」という結果から「仮説は○」と判断しても大外れでない</u>
といえる
なぜならば　前に説明したように
θ_0 近く($\theta \fallingdotseq \theta_0$)では検出力は α 程度しかない
つまり本当は仮説は×なのに「×でない」と判断する確率が高いが　そもそも $\theta \fallingdotseq \theta_0$ だから「仮説は○(=真の値は　θ_0)」だと判断しても大きな誤りにはならない
<u>*「当たらずとも遠からず」*</u>!
逆にいえば
θ_0 の近くで B(θ)が急激に凹にならない場合には
<u>「仮説が×でない」という結果から「仮説が○」だと判断してはいけない</u>
例えばサンプル規模 n が小さい場合など
実際の応用ではこのちがいはとても重要
⇒検出力関数の形状を頭にいれておくと　いろんなことが明確に整理できる

正規分布表

a に対して、1-Φ(a)を与える。但し、Φ(a)は標準正規分布の分布関数である。
すなわち、Z~N(0,1)のとき、Φ(a)=P(Z≦a)

a	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500	0.496	0.492	0.488	0.484	0.480	0.476	0.472	0.468	0.464
0.1	0.460	0.456	0.452	0.448	0.444	0.440	0.436	0.433	0.429	0.425
0.2	0.421	0.417	0.413	0.409	0.405	0.401	0.397	0.394	0.390	0.386
0.3	0.382	0.378	0.374	0.371	0.367	0.363	0.359	0.356	0.352	0.348
0.4	0.345	0.341	0.337	0.334	0.330	0.326	0.323	0.319	0.316	0.312
0.5	0.309	0.305	0.302	0.298	0.295	0.291	0.288	0.284	0.281	0.278
0.6	0.274	0.271	0.268	0.264	0.261	0.258	0.255	0.251	0.248	0.245
0.7	0.242	0.239	0.236	0.233	0.230	0.227	0.224	0.221	0.218	0.215
0.8	0.212	0.209	0.206	0.203	0.200	0.198	0.195	0.192	0.189	0.187
0.9	0.184	0.181	0.179	0.176	0.174	0.171	0.169	0.166	0.164	0.161
1.0	0.159	0.156	0.154	0.152	0.149	0.147	0.145	0.142	0.140	0.138
1.1	0.136	0.133	0.131	0.129	0.127	0.125	0.123	0.121	0.119	0.117
1.2	0.115	0.113	0.111	0.109	0.107	0.106	0.104	0.102	0.100	0.099
1.3	0.097	0.095	0.093	0.092	0.090	0.089	0.087	0.085	0.084	0.082
1.4	0.081	0.079	0.078	0.076	0.075	0.074	0.072	0.071	0.069	0.068
1.5	0.067	0.066	0.064	0.063	0.062	0.061	0.059	0.058	0.057	0.056
1.6	0.055	0.054	0.053	0.052	0.051	0.049	0.048	0.047	0.046	0.046
1.7	0.045	0.044	0.043	0.042	0.041	0.040	0.039	0.038	0.038	0.037
1.8	0.036	0.035	0.034	0.034	0.033	0.032	0.031	0.031	0.030	0.029
1.9	0.029	0.028	0.027	0.027	0.026	0.026	0.025	0.024	0.024	0.023
2.0	0.023	0.022	0.022	0.021	0.021	0.020	0.020	0.019	0.019	0.018
2.1	0.018	0.017	0.017	0.017	0.016	0.016	0.015	0.015	0.015	0.014
2.2	0.014	0.014	0.013	0.013	0.013	0.012	0.012	0.012	0.011	0.011
2.3	0.011	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.009	0.008
2.4	0.008	0.008	0.008	0.008	0.007	0.007	0.007	0.007	0.007	0.006
2.5	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.005
2.6	0.005	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
2.7	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
2.8	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
2.9	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001
3.0	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

t 分布表

t が自由度 m の t 分布に従うとしたとき、t の上側 100α%点を与える。
すなわち、$P(t \geqq u)=\alpha$ となる u を与える。ここで $u=t_\alpha(m)$である。

自由度m ＼ α	0.1	0.05	0.025	0.01
1	3.078	6.314	12.706	31.821
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.896
9	1.383	1.833	2.262	2.821
10	1.372	1.812	2.228	2.764
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602
16	1.337	1.746	2.120	2.583
17	1.333	1.740	2.110	2.567
18	1.330	1.734	2.101	2.552
19	1.328	1.729	2.093	2.539
20	1.325	1.725	2.086	2.528
21	1.323	1.721	2.080	2.518
22	1.321	1.717	2.074	2.508
23	1.319	1.714	2.069	2.500
24	1.318	1.711	2.064	2.492
25	1.316	1.708	2.060	2.485
26	1.315	1.706	2.056	2.479
27	1.314	1.703	2.052	2.473
28	1.313	1.701	2.048	2.467
29	1.311	1.699	2.045	2.462
30	1.310	1.697	2.042	2.457

カイ２乗分布表

Y が自由度 m のカイ 2 乗分布に従うとしたとき、Y の上側 100α%点を与える。
すなわち、P(Y≧u)=α となる u を与える。ここで、u は $\chi^2_\alpha(m)$に等しい。

α / 自由度m	0.99	0.975	0.95	0.05	0.025	0.01
1	0.00016	0.00098	0.0039	3.84	5.02	6.63
2	0.020	0.051	0.10	5.99	7.38	9.21
3	0.11	0.22	0.35	7.81	9.35	11.34
4	0.30	0.48	0.71	9.49	11.14	13.28
5	0.55	0.83	1.15	11.07	12.83	15.09
6	0.87	1.24	1.64	12.59	14.45	16.81
7	1.24	1.69	2.17	14.07	16.01	18.48
8	1.65	2.18	2.73	15.51	17.53	20.09
9	2.09	2.70	3.33	16.92	19.02	21.67
10	2.56	3.25	3.94	18.31	20.48	23.21
11	3.05	3.82	4.57	19.68	21.92	24.72
12	3.57	4.40	5.23	21.03	23.34	26.22
13	4.11	5.01	5.89	22.36	24.74	27.69
14	4.66	5.63	6.57	23.68	26.12	29.14
15	5.23	6.26	7.26	25.00	27.49	30.58
16	5.81	6.91	7.96	26.30	28.85	32.00
17	6.41	7.56	8.67	27.59	30.19	33.41
18	7.01	8.23	9.39	28.87	31.53	34.81
19	7.63	8.91	10.12	30.14	32.85	36.19
20	8.26	9.59	10.85	31.41	34.17	37.57
21	8.90	10.28	11.59	32.67	35.48	38.93
22	9.54	10.98	12.34	33.92	36.78	40.29
23	10.20	11.69	13.09	35.17	38.08	41.64
24	10.86	12.40	13.85	36.42	39.36	42.98
25	11.52	13.12	14.61	37.65	40.65	44.31
26	12.20	13.84	15.38	38.89	41.92	45.64
27	12.88	14.57	16.15	40.11	43.19	46.96
28	13.56	15.31	16.93	41.34	44.46	48.28
29	14.26	16.05	17.71	42.56	45.72	49.59
30	14.95	16.79	18.49	43.77	46.98	50.89