

数理統計学 Q & A (2019)

数理統計学に関連する質問

(Q) 確率・統計の歴史的側面を書いた本でおススメのものはありますか？

(A) 私は科学史を学問として研究したことはないが、歴史は好きです。

[1] イアン・ハッキング(石原英樹・重田園江訳)「偶然を飼いなす」木鐸社, 1999.

「この博物誌的な書物を好奇心に満ちたすべての読者に捧げる」とある。確率統計が 20 世紀の科学に中でいかに成功してきたかを科学的な視点で論ずる。かなり興味深い。

[2] 西内啓「統計学が最強の学問である」ダイヤモンド社。

統計リテラシーの必要性をジャーナリスティックに説く話題の本。血が沸き立つような書きぶりではあるが、フィッシャー(20 世紀前半の大統計学者)を超えるのは難しい。

[3] イアン・ハッキング (広田すみれ・森元良太訳)：確率の出現，慶應義塾大学出版会，2013.

[4] キース・デブリン (原 啓介訳)：世界を変えた手紙 --- パスカル、フェルマーと〈確率〉の誕生，岩波書店，2010.

この 2 冊は確率論の始まりをさまざまなエピソードとともに語る。個人的にはカルダーノ(1501--1576)という人物に大変興味がある。

[5] 小杉 肇：統計学史，恒星社厚生閣，1984.

大部なだけあって、色々なことが書いてあって役に立つ。

(Q) 問題だと有意水準を自分で決めてから、実現値を使って計算して仮説検定をしていますが、その過程は逆ではダメなんでしょうか？例えば、ある確率を異常値だと言い張りたい人がいたら、実現値を使った計算をして、それがぎりぎり異常値になるように有意水準を調整しそうだと思いました。

(A) 本来、有意水準は数理統計学ではなく、外的要因（リスクとかコストとか）によってきめてかかるものです。有意水準から決まる棄却域に実現値が落ちるかどうかで意思決定をするわけです。言いたいことは理解できます。その文脈では、実現値がどれほど稀かを与える P 値と呼ばれる確率を示すことが正しいやり方です。棄却なり採択なりの結論ありきで、有意水準を定めたのでは、あたかも統計的な保証があって意思決定をしたかのような偽装になってしまいます。薬剤の効果の評価などに対してこのような疑いが持たれがちです。

(Q) 紅茶は茶葉から煎れると全く香りが違うのでぜひやってみてください。

(A) フィッシャーは、自身が構築した仮説検定のアイデアをやさしく説明するために、ミルクティーの話を使った。ミルクティーは英国ではとてもポピュラーな飲み物ですね。その

ミルクティーを入れるときに、紅茶が先かミルクが先かによって味が変わるという。その違いが判るといふご婦人に対して、実際に判別実験をする。その結果から、このご婦人が本当に味の違いを分かっているのか、実はランダムに答えているのかを統計的に判定するという話です。

(Q) 二項母集団の母比率 p を推定するために、

$$Z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \sim N(0,1)$$

であることを用いて、さらに大数の法則から $\hat{p} \approx p$ として

$$Z = \frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}} \sim N(0,1)$$

となることが用いられました。この方法で、実際に母比率が推定できることは分かったのですが、そもそも \hat{p} から p を推定するというプロセスの中で $\hat{p} \approx p$ を適用することに違和感を感じてしまいました。しかもそれを分母にだけ適用しているので少し腑に落ちませんでした。

(A) はい、なかなか鋭い指摘ですね。 $\hat{p} \approx p$ を用いる前の

$$Z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \sim N(0,1)$$

からでも同じ信頼区間を出すことができます。例えば 95% 信頼区間であれば、

$$-1.96 \leq Z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \leq 1.96$$

とおいて、この式を p について解けばよいのです。2 次不等式になるので具体的に（式は長くなっても）解くことができます。その式には平方根が含まれるので、そこを 1 次近似することで、

$$\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

が導出されます（教科書参照）。この結果は、大数の法則を用いた議論と結果的に同じものになります。大数の法則を用いた議論は、ご指摘のとおりやや厳密さに欠けますが、わかりやすいので使いました。

(Q) 信頼係数を 90%や 95%にするのはなぜですか？

(A) 習慣です。皆が同じ信頼度を使えば、比較も容易になりますし。起源は、フィッシャーが理論を作ったとき、95%を例示としてよく使ったためだと思われます。95%の信頼度というのは、20 回に 1 回は外すが、19 回は当てるといふくらいの確実さでモノを言うということになります。フィッシャーは農業への応用も念頭にあったそうで、毎年のことを推定す

るのなら、20年に1回外すくらいで丁度よいと考えたらしいです。なにせ100%にはできないので。

(Q) 今まで自分が使っていた標本平均に加えて、母平均の区間推定という評価の方法があることが分かった。どのくらいの範囲の中に母平均が含まれているかを推定できることの有用性がよくわかりためになった。

(A) 近年では、算術平均だけで誤差評価のない結果は論文として評価されません。区間推定は様々な文脈で使われる重要な処理法なのでよく学習してください。

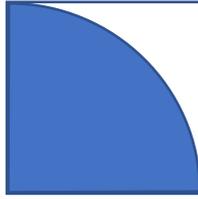
(Q) 今回の授業とは話が変わるのですが、宝くじの期待値は調べたところだいたい0.5でした。もし仮に宝くじを買うとしたら、少ない枚数を買うのとたくさんの枚数を買うのだとどっちのほうが得とみなすことができますか？たくさん買ったら、期待値通りの結果になると思うし、少なく買ったら必ずしも期待値通りの結果にはならないし、気になりました。

(A) 「宝くじの期待値」が何をさしているのかわかりません。宝くじ(1枚100円)を1枚買った時の賞金の期待値なら35円くらいではなかったでしょうか？調べればはっきりしますが、仮に35円、つまり回収率35%として話を進めます。宝くじを大量にランダムに買えば、大数の法則が効いてきて、賞金は投資額の35%に近づくことでしょう。しかも、中心極限定理が効いてきて、35%の前後の揺らぎはどんどん小さくなり、おおむね35%が回収されることになるでしょう。したがって、宝くじを毎週買い続ける人は、大数の法則を実践することになり、回収率は35%にどんどん近づくということです。おススメできません。逆に、買う枚数が少なければ、平均的には35%回収できることに変わりはないですが、揺らぎが大きいので、大勝ちすることもまだまだそれなりに期待できるということです。さらに、「勝ち逃げ」が戦略としてはよいこともこれからわかります。ただし、これは宝くじをランダムに買う場合です。シリーズで買うなどすると当然確率計算は変わります。

(Q) 演奏会の運営を担当した時に、チケットの前売り券を渡しすぎて、当日、座席数を上回る来場者数となりトラブルになった。今回、話に合ったような確率を考えてリスクをコントロールできるなと感じたし、面白いなと思いました。

(A) ぜひ、実践に役立ててください。何%位が当日来ないかが経験的にわかれば、どのくらい余分に売ると、当日どのくらいの確率でトラブルになるかがわかります。

(Q) 中学校の数学の授業で、左のような(大きく)正方形と四分円を描いたエリア内に、大量に点を打ち、円内の点を数えてみるという実験をしたことがあります。円の中に点が打たれる確率は



$$\frac{\text{四分円の面積}}{\text{正方形の面積}} = \frac{\pi}{4}$$

となるという話を聞いたので、面積で考えることが確かにできるなあと納得したのを思い出した。(でも、その時は、円周率が自分の手計算でわかるということに感動しました。)

(A) 中学校のときにこんなことを確認しているのですね。まさに、モンテカルロ法の実証ですね。どのくらいの精度が出たのか気になります。

(Q) 確率は $0 \leq p \leq 1$ ですが、同じように考えると%表示も 0% から 100% までしか表せないと思います。ですが、ネットやスーパーでは 150%増量とかよく聞きます。これは使い方として正しいのでしょうか？

(A) はい、正しいです。もとの量を 100%として 150%増量されて 250%になったということです。つまり、量が 2.5 倍になったのですね。一方、確率は 200%とかダメです。かつて、選挙前に有力候補が「私が立候補することは 200%ない」とか言っていました。日常語の中の確率という観点からは、「私が立候補することは 100%ない」と言えばよいのです。200%などと言うと、強調していることは伝わりますが、学がないと思われる。別の話、昔、降水確率が使われ始めた頃、「土曜日が晴れる確率は 80%、日曜日が晴れる確率も 80%です。なので、週末が晴れる確率は 160% となって、雲一つない快晴になるでしょう」などと言っていた人がいたらしい。

(Q) 結果から原因が探れるというベイズの公式の応用例が興味深いです。

(Q) 条件付確率には苦手意識があったがけど、今日でなくなりました。

(A) 条件付確率は直感に合わないこともあるので、使い方は要注意ですが、たいへん面白い。

雑談

(Q) 僕の出身の千葉県でチェーン展開している「ちばチャン」という店のから揚げのぼか盛りというのが、大人 5 人でやっと食べれるかという量なのに 1500 円くらいで、しかも本当においしいので機会があったら大人数で言ってきてください。1 人だと絶対に無理です。

(A) 1 人用のから揚げでよいです。から揚げは好きなので、是非行ってみたいです。

(Q) マックのポテト 150 円セールに行ったら、店内の人のほとんどがポテトしか頼んでなくて少し異様でした。先生も行ってみてください。

(A) マックにはあまり行かないなー。ある時期からコーヒーがそれなりになったので、コーヒーだけ頼むことはありました。ポテトだけとはちょっと様にならない気がする。

(Q) 最近、仙台駅近くの朝市でお魚を仕入れて、自分でさばくのにはまっています。ストレス発散になるので、時間があるときやってみてください。ちなみに最近では、アジの開きを作りました。

(A) すばらしい趣味をお持ちですね。いつも美味しい魚が食べれそうです。魚をさばくのは、私にはかえってストレスになりそうです。

(Q) わたし、ギネス大好きなんです。

(A) ぜひ、現地に赴いて堪能してください。博物館にゴセットのコーナーがあるみたいなので、そこも見てください。

(Q) そういえば、私は鳥人間コンテストに出場する Windnauts という部活に所属しています。8 月末、テレビ放映があるはずなので、ぜひ見てくださいね！最近部活が忙しくて、授業中寝てしまうことが多いです。本当にすみません。

(A) 本学の Windnauts は常連ですよ。最近、残念ながら見ていません。私が若いころに始まったコンテストです。かつてはちょっと飛ぶだけだったのが、いつのまにか琵琶湖を往復するくらいになって、技術革新ぶりに驚いたものです。どのような役どころでしょうか？ともかく、頑張って東北大学の名声を高めてください。

(Q) この前、長町の「びすたーり」というところに行き、シソの利いたシジミのスープを飲んだのですが、シソもシジミもあまり好きではなかったにも関わらず、ほど良いバランスになっていておいしく感じたので、ぜひ飲んでみてください。

(A) おいしそうです。長町は（ほぼ）地元なので「びすたーり」知ってますよ。あの店の最大の売りはベーゼンドルファーが置いてあることです。見せてもらい、低音部の黒健から察するに、2 千万円近いモデルかなと推定しました。子供も弾いているそうですが。

(Q) 地元青森にある「あさ利」というラーメン屋のねぎラーメンが超辛くておいしいので、青森に行く機会があったら行ってみてください。

(A) はい、ぜひ。青森は遠いけど。

(Q) B 級映画レポート「シャークネード カテゴリー 2」前回紹介した「シャークネード」の続編である。前作に続き、あいかわらずの B 級ぶりである。前作でサメに打ち勝った主

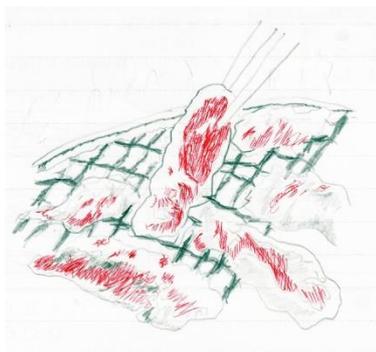
人公。離れて暮らす妻の家族に会うため NY を訪れるが、そこにまたサメを巻き上げた竜巻がやってくる。サメとの戦いがまた始まる。(以下略)

(A) こういうの好きだねー。熱い語りの中に、B 級映画への愛を感じました。サメ映画だと「ロストバケーション」が、海の映像がきれいでよかったです。ちょっとエグいけど。

(Q) 昨日、部活の新歓後に自販機前で現役と 1 年生との間で自己紹介をした際に、「好きなテレビ番組は？」と言うのを 1 人ずつ言ってゆく流れになったところ、13 人いたうちのほとんどが「最近 TV 見ないのですが。。。」から始まり、、、(中略) しっかり TV を見ていたのは嵐が好きだという現役部員 1 人くらいでした。

(A) 私自身も TV なしの生活が 15 年以上になるので観る機会がかなり少ないです。が、僕たちの世代はテレビっ子で育った人が多いので、大学の高齢教師は TV 見ている人が多いかも。巷間、言われることは、最近の TV のメインターゲットは高齢者ということですから、その通りなんでしょうね。

(Q)



(A) 講義を聞きましょう。

(Q) 春休みにグレゴリー・ホブリット監督の「オーロラの彼方に」を観ました。一連の殺人事件に過去と未来から迫る SF です。「おまえの 30 年をなかったことにしてやる」(劇中より)

(A) 未見。黒歴史がなくなるのならそれに越したことはない。。。

(Q) ラプラスの悪魔という話を思い出した。アカシックレコード理論とも通じるものがあると思う。ちなみに、ラプラスの悪魔をネタにしたライトノベルもあります。

(A) アカシックレコード理論とは切り離れていると思う。ラプラスの悪魔ネタといえば「ラプラスの魔女」という小説があり、映画にもなった。この小説は学生さんに教えてもらい、さっそく本を読んで私も観た。ま、そういうことね。

(Q) 私のアルバイト先「●●●」の店主が Fisher に似ています！ (中略) 先生が蕎麦嫌いでないことが大前提ですが、気が向いたらぜひお越しください。

(A) 興味津々です。たまに、山形まで蕎麦を食べに行ってますよ。ぜひ機会を作って、のぞきにまいります。

(Q) Novo Amor という歌手がおすすめです。とてもリラックスできます。半年間よろしくお願いします。

(A) 情報をありがとうございます。

(Q) オーストラリアのエアーズロックが今年の 10 月から登れなくなるので、それまでに行った方が良さそうです。

(A) オーストラリアには古い友人がおり、ごくたまに訪ねてます。シドニー、メルボルン、ベンディゴだけですが。シドニーとか、目下、物価がとても高いが、住めるのなら良いところだなあと感じます。友人は毎朝、海辺のカフェで朝コーヒーしたりしています。エアーズロック、ぜひ見てみたいと思いますが、実現するかなあ。。。

(Q) ペンギンの足がしもやけにならない理由 (以下省略)

(A) 豆知識ですね。

(Q) 「予備校のノリで学ぶ大学数学・物理」という YouTube のチャンネルが面白い。ベイズの定理の動画もあります。去年、このチャンネルのおかげで線形代数が浮きました。

(A) YouTube はじめネット上には面白い講義が山のようにありますね。中には、ほんとうに優れたものもあります。凡百の大学教師が教室で講義するのはかえって害かも知れません。