

飼い猫は社会空間認識能力を持つか？

対応のない 2 標本平均値の差の検定：シミュレーションに基づくアプローチ

学習目標

- 2つのカテゴリーを持つ説明変数について、「関連がない」対「関連がある」の観点から帰無仮説と対立仮説を、平均値（すなわち、 μ_1 と μ_2 ）の比較の観点からも述べる。
- 2標本の平均を比較するために 3S 戦略を実行する：(I: Statistic) 統計量を求め、(II: Simulation) シミュレーションを行い、(III: Strength of Evidence) 観察された研究結果が偶然だけで起こったものでないことの証拠の強さを計算して示す。
- 2つの変数の間に関連がない場合、繰り返される無作為割付けでどのような結果（平均値または中央値の差）が期待されるかをシミュレーションするためにカードを使用する方法を採用する。
- 帰無仮説のシミュレーションを行うために [Multiple Means アプレット](#) を使用し、アプレットからの出力を理解する。
- 2標本の平均値の差の検定の標準化統計量と p-値を求め、解釈する。
- 2SD 法を用いて、2つのグループ間の母平均の差の 95%信頼区間を求め、その区間を研究の文脈で解釈する。平均値の差の 95%信頼区間がゼロを含むことの意味を解釈する。
- p-値および／または標準化統計量と、統計的有意性、推定、一般化可能性、因果関係を含む研究デザインに基づいて、対立仮説（および帰無仮説）について完全な結論を述べる。

ステップ 1：リサーチ・クエスチョンを形成する

多くの野生の動物にとって、実際に動物を見ていなくても他の動物を追跡できることは重要である。動物によっては、音やにおいといった他の感覚を使ってこれを行うものもいる。他の動物の所在を心の中で想い続けることによって、その動物を追跡することは、社会空間認識の一形態である。日本の研究者たち（高木ら、2021 年）は、一般的な飼い猫が社会空間認識を持つかどうかを調べようとした。



これらの資料は STUB ネットワークによって開発され、NSF- DBI 1730668 助成金の下、全米科学財団の支援を受けています。これらの資料は、クリエイティブ・コモンズ・ライセンス BY-NC のもと、非営利目的に限り、また STUB ネットワークへの帰属が明記されている場合に限り、利用者が資料を配布、翻案、構築することを許可されています。

ステップ2：研究を計画し、データを収集する

研究者たちは、猫に部屋の外で飼い主の声を聞き、部屋の反対側で突然その声を聞くという「テレポーテーションのような」状況を与えた。もし猫に社会空間認識能力があれば、飼い主の位置が突然変わったように見えて、猫は驚きを示すだろうと研究者たちは考えた。研究者たちは40匹の猫に、自宅か猫カフェのような慣れ親しんだ環境でテストを行った。最初のスピーカーは、猫がいる部屋のドアの反対側に置いた。2つ目のスピーカーは、部屋の反対側、1つ目のスピーカーから4メートル以上離れたドアか窓の近くに置いた。猫は最初のスピーカーから飼い主の声を聞く。そして2.5秒後、2番目のスピーカーから飼い主の声か別の人の声が聞こえる。2番目のスピーカーからの音声を聞いた後、8人が猫が示した驚きの度合い（耳をずらす、前後を見る、動く）を0点（驚かない）から4点（強く驚く）までで評価した。評価者のスコアを平均し、各猫に驚きスコアを与えた。もし一般的な飼い猫が社会空間認識能力を持っているならば、2番目のスピーカーから聞こえてくる飼い主の声を聞いた猫は、違う声を聞いた猫よりも高いレベルの驚きを示すだろうと研究者たちは考えた。40匹の猫を無作為に2つの条件に割り当てた。

1. 観察単位（または実験単位）は何か？
2. この研究における説明変数と目的変数を特定せよ。また、それらを質的変数か量的変数かに分類せよ。
3. これは実験か、もしくは観察研究か？どのように判断したのか説明せよ。
4. 2つ目のスピーカーで同じ声（飼い主の声）を聞いた猫の驚きスコアの母集団平均を μ_{same} とし、2つ目のスピーカーで違う声を聞いた猫の驚きスコアの母集団平均を $\mu_{\text{different}}$ とする。このとき、同じ声を聞くことが猫の驚き度に影響を与えるかどうかを調べるための帰無仮説と対立仮説を言葉と記号を用いて述べよ。注: 猫が別の声を聞くことでより高いレベルの驚きを示す可能性があるため、両側検定を用いる。

ステップ 3: データを調べる

5. **Surprise** データを **Multiple Means アプレット** に入力し、**データを使う** を押す。目的変数として **Surprise** を選択し、説明変数として **Voice** を選択する。

a. アプレットは、各研究グループに 1 つずつ、平行なドットプロットを作成することに注意する。これらのグラフに箱ひげ図を追加することも可能である。これらのグラフだけに基づいて、どのグループ (same か diff か) の平均が、より高い驚きスコアを持っているように見るか? 考えを答えるだけでなく、どのように判断したのかも説明せよ。

b. このアプレットは、各群の驚きスコアの平均や標準偏差 (SD) のようなデータの数値要約も計算することに注意する。

i. same グループについて、標本数 (n)、平均、標準偏差(SD) を記録する。

$$n_{\text{same}} = \quad \bar{x}_{\text{same}} = \quad SD_{\text{same}} =$$

ii. diff グループについて、標本数 (n)、平均、標準偏差(SD)を記録する。

$$n_{\text{different}} = \quad \bar{x}_{\text{different}} = \quad SD_{\text{different}} =$$

標準偏差は変動性の尺度である。相対的に言えば、標準偏差の値が小さいほどばらつきが小さく、標準偏差が大きい分布と比較して、データ値がより密接に集まりやすい分布であることを示す。

c. 数値の要約に基づき、どちらのグループ (same グループまたは diff グループ) の驚きスコアの平均が高かったか述べよ。また、(a)の予測は正しかったか? 驚きスコアのばらつきが大きかったのはどちらのグループか? についても答えよ。

d. このアプレットは、2つのグループ間で観察された驚きスコアの平均の差も報告していることに注意して、この値を記録する。推測分析を行う前に、標本平均のこの差は意味のある差だと思うか? 考えを答えるだけでなく、そのように答えた理由も説明せよ。

$$\bar{x}_{\text{same}} - \bar{x}_{\text{different}} =$$

ステップ4：データを超えた推論を導く

6. 2つのグループの驚きスコアの平均が異なる値となった理由について、考えられる説明を2つ挙げよ。

観察された結果が、同じ声を聞くことが猫の驚き度に影響を与えるという説得力のある証拠となるかどうかを判断するために、シミュレーション分析を使って、同じ声を聞くことが猫の驚き度に影響を与えないと仮定した場合に得られる結果に対して、得られた結果が典型的なものなのか異常なものなのかを判断せよ。

シミュレーション分析のポイントは、もし同じ声を聞いても驚き度に影響がないのであれば（帰無仮説）、同じ声を聞いても聞かなくても猫の反応は変わらなかったと仮定することである。言い換えれば、このシミュレーションは帰無仮説が真であること、つまり同じ声を聞いたか否かと驚き度との間に関連性がないことを前提としている。

キーポイント

観察された2つのグループ間の差の統計的有意性を評価するために、説明変数グループについて目的変数の値の無作為割付けから、標本平均の差がどのような傾向にあるかを調査する。

7. 上記のキーポイントに当てはめて、観察された2つの標本平均の差が統計的に有意かどうかを判断する方法を説明せよ。具体的には、どのように same と diff の各グループ内データについて目的変数の値の無作為割付けを実施するか説明せよ。

繰り返しになるが、重要な問題は、もし本当に同じ音声は驚きスコアに効果がなかったとしたら、無作為割付けだけで、この研究で観察された差よりも極端な群間差がどれくらいの頻度で生じるかということである。このような極端な差が偶然だけで（差がない／効果がない／関連がないという帰無仮説が真であった場合）どれくらいの頻度で生じるかを調べるためにシミュレーションを用いることができる。再び3S戦略を採用せよ。

I. (Statistic) 統計量

8. 観察されたグループの平均が互いにどの程度異なるかを測定するための自然な統計量は、2つのグループの間の驚きスコアの平均の差である。#5(d)で報告したように、この統計量の値を報告せよ。

II. (Simulation) シミュレーション

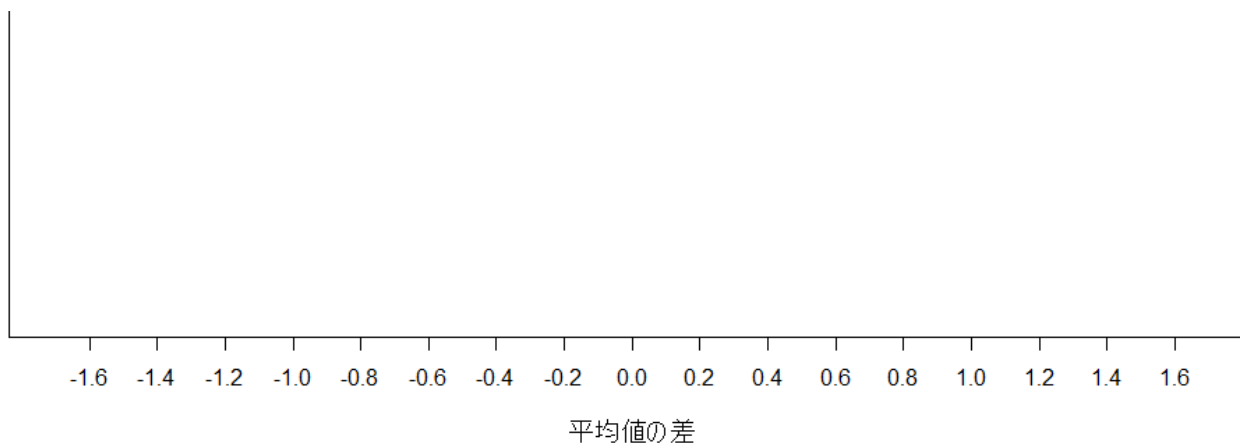
まず、声の状態が驚きスコアに影響しないと仮定して、40 匹の猫（とその驚きスコア）を 2 つのグループにランダムに割り当てるシミュレーションを、インデックスカードを使って行うことから始める。帰無仮説では、驚きスコアは声の状態とは無関係であるため、40 匹の猫がどちらの声の状態のグループ（同じグループ、異なるグループ）に振り分けられたとしても、驚きスコアは全く同じであったと仮定する。

9. このシミュレーションを行うには:

- a. カードは何枚必要か？
- b. それぞれのカードに何を書くか？
- c. このシミュレーションを 1 回繰り返すには、40 枚のカードの束をよくシャッフルし、カードをランダムに 2 つのグループに分ける。一方のグループには 21 枚のカード（同じ声）があり、もう一方のグループには 19 枚のカード（違う声）がある。
 - i. 再ランダム化された各グループの標本平均を計算して報告する。
 - ii. (i)で報告した平均の差を計算する：つまり、同じ声の平均から異なる声の平均を引いた値を報告する。

なお、上記の操作は、エクセルファイルでも同様に操作できる。

- d. この結果とクラスメートの結果を組み合わせ、声の状態が驚きのスコアに影響を与えないと仮定した場合の、標本平均の差のいくつかの可能な値の分布を示すドットプロットを作成する。ドットプロットでは、横軸にラベルを付ける。



- e. ドットプロットの中心はどの程度の値か？なぜそうなるのか説明せよ。
(ヒント：シミュレーションを行うとき、何を仮定しているか？)
- f. 観察された平均値の差（#8 で報告されたもの）は、ドットプロットのどこにあるか？この値は頻繁に起こったか、やや稀に起こったか、それともめったに起こらなかったか？どのように判断したのかも説明せよ。
10. 声の状態が驚きスコアに影響しないと仮定して、グループの各データの平均値の差の典型的なものと典型的でないものを決定するために、さらに何度も何度も繰り返し実施する。手作業で長時間カードをシャッフルして群平均の差を計算するよりも、コンピュータを使う方がより好ましい。[Multiple Means アプレット](#)に戻り、シャッフル・オプションの表示ボックスをチェックし、プロット表示を選択し、目的変数をシャッフルを押す。
- a. アプレットが何をしているのか、そしてアプレットの操作が#4 の帰無仮説とどのように関係しているのかを説明せよ。
- b. アプレットの出力で与えられるように、再ランダム化された群の標本平均のシャッフルされた差を記録せよ。この差は（#8 で報告された）研究で観察された差よりも極端であるか？考えを答えるだけでなく、どのように判断したのかも説明せよ。
- c. [Multiple Means アプレット](#)に戻り、もう一度 目的変数をシャッフルをクリックして、再ランダム化されたグループの標本平均の差を記録せよ。#10(b)の結果と変わったか？
- d. もう一度 目的変数をシャッフルをクリックして、再ランダム化されたグループの標本平均の差を記録せよ。#10(b)と#10(c)の結果と変わっただろうか？

ここで、音声条件がサプライズスコアに影響を与えないと仮定して、標本平均の差の多くの可能な値を見るために、[Multiple Means アプレット](#)で次のようにする。

- シャッフルの数を 1 から 997 に変更する。
- 目的変数をシャッフルを押して、合計 1,000 回のシャッフルと統計量の再ランダム化を行う。

- e. 標本平均値の差の 1,000 のシミュレーション値のドットプロットを考察する。
- ドットプロットの 1 つの点は何を表しているか？(ヒント: グラフ上にもう一つ点を追加するには何をしなければならないかを考えてみる。)
 - このドットプロットに表示されている帰無分布の全体的な形を記述せよ。
 - 観察された標本平均の差は (#8 で報告されたように) このドットプロットのどこに位置するか？音声条件が驚きに影響を与えないと仮定した場合、観察された差よりもさらに極端な点がたくさんあるといえるか？また、どのように判断したのか説明せよ。
- f. p-値を推定するために、再び [Multiple Means アプレット](#) を使用する。研究で観察された same と diff の 2 グループ間の平均の差 (#8 で報告されたもの) を入力し、シャッフルを数えるの横で適切な対立仮説を選び、数えるを押す。出力に基づくおよその p-値は何か？
- g. p-値の解釈を述べよ。

帰無仮説 (同じ声を聞いても驚きスコアに影響がない) が真であった場合、観測された統計値 (= _____), あるいはそれより極端に帰無仮説より統計値が離れる確率は _____。

III. (Strength of Evidence) 証拠の強さ

11. p-値に基づいて、音声条件は驚きスコアに影響を与えないという帰無仮説に対して、データによって提供された証拠の強さを評価する：「あまり証拠がない、中程度の証拠がある、強い証拠がある、または非常に強い証拠がある。」のうち、適切な証拠の強を選ぶ。
12. 2SD 法を用いて、同じ声を聞いた猫の平均驚きスコアから、違う声を聞いた猫の平均驚きスコアを引いた差の 95%信頼区間を概算する。(ヒント：研究で観察された same と diff の 2 グループ間の平均の差とシャッフルされた same と diff の 2 グループ間の平均の差の SD の推定値は、アプレットのシミュレーション結果から生成できる。95%信頼区間 = 平均値の差の観察値 \pm 2SD)。

ステップ5：結論の策定

13. **意義**：この研究の文脈における証拠の強さに関して、結論を要約する。

（論文では、研究者たちはより大規模なデータセットを分析した。その分析結果に基づいて研究者たちは、猫が社会空間認識能力を持っている可能性を示していると述べた。）

14. **推定**：この信頼区間が何を明らかにするかについて、特に区間が完全に正であるか、完全に負であるか、またはゼロを含むかに注意して、解釈を述べよ。

15. **一般化**：この研究の猫はより大きな集団から無作為に選ばれたのか？この研究の結果をどのような母集団に一般化してもよいのか？

16. **因果関係**：研究参加者は音声条件に無作為に割り当てられたのか？このことは結論にどのような影響を与えるのか？

ステップ6：振り返り、先を見据える

17. **振り返り**：この研究のデザインと結論について、何か気になる点はあったか？特に、研究で観察された same と diff の2グループ間に真の差があることを示す強力な証拠を発見する可能性を高めるためにできたことはあるか？批評したい事項は以下の通り：

- 研究課題と研究デザインのミスマッチ
- 観察単位を選択方法
- 観察単位にどのように処置が割り当てられたか
- 測定値の記録方法
- 研究における観察単位の数
- 観察されたことが実用的価値があるかどうか

18. **先を見据える**：限界を修正し、この知識を構築するために、研究者たちは次のステップに進むべきか？

さらに探る：もうひとつの統計量

19. **Multiple Means アプレット**に戻り、**統計量**プルダウンメニューを使って**中央値の差**を選択する。

- a. 元データの**要約統計量**から、各グループの驚きスコアの中央値を記録する。また、同じ声の中央値から異なる声の中央値を引いた値を用いて、中央値の差も記録する。
- b. **シャッフルの数**に 1000 を入力し、**目的変数をシャッフル**を押す。グループの中央値の差の結果の帰無分布を記述する。この帰無分布は、ゼロ付近に中心があるように見えるか？釣鐘型の分布に見えるか？
- c. p-値を推定するには、観察されたグループの中央値の差を入力し、**シャッフルを数える**の隣にある適切な対立仮説を選択して、**数える**を押す。出力に基づいて、観察された中央値の差とおおよその p-値を報告する。
- d. この p-値は、驚きスコアに対する音声効果の強い証拠を示しているか？考えを答えるだけでなく、どのように判断したのかも説明せよ。
- e. この研究において、どちらの統計量（平均値の差と中央値の差）の方が、驚きのスコアに音声効果があることを示すより強い証拠が得られるか？考えを答えるだけでなく、どのように判断したのかも説明せよ。
- f. 次の探査では、研究で観察された same と diff の 2 グループの平均の間に差の強い証拠があるかどうかを決定するために、理論に基づく検定を使用する。パート (b) で見た中央値の差の帰無分布に基づき、平均値の代わりに中央値を使うことが理論に基づく検定で問題になる理由を説明せよ。

参考

Takagi S, Chijiwa H, Arahori M, Saito A, Fujita K, Kuroshima H (2021) Socio-spatial cognition in cats: Mentally mapping owner's location from voice. PLoS ONE 16(11): e0257611.
<https://doi.org/10.1371/journal.pone.0257611>

見えない飼い主を耳だけで追跡 猫の空間認識能力、実験で実証
<https://www.cnn.co.jp/fringe/35179307.html>

ネコはいつでも飼い主を思っていた...常に居場所を頭の中で追跡との調査結果が
https://www.newsweekjapan.jp/stories/world/2021/11/post-97450_1.php

講義内容は STUB ネットワークのメンバーによって独自に作成されたものです。日本語翻訳チーム：Jimmy Doi, Cal Poly San Luis Obispo (jdoi@calpoly.edu)、橋本三嗣 広島大学附属中・高等学校 (mhashimo@hiroshima-u.ac.jp)、中島康彦 群馬県立前橋高校 (nakajima-yshk@edu-g.gsn.ed.jp)。翻訳に関するご質問は翻訳チームまでお願いします。