

問題

下記の小問について解答を submit.yml に記載して提出してください。

小問 1

以下の選択肢の中から正しい選択肢をすべて答えてください。

1. MAP@K はレコメンドシステムのランキング評価を行うための一つの指標である
2. 時系列データの予測モデルに対してバリデーションを行う際に、期間を考慮しない hold-out 法を利用する
3. Adversarial Validation を用いて特徴量選択を行う場合、学習データとテストデータを分類するような分類器を作成し、その分類器の性能を下げるような特徴量を選択する
4. 重回帰モデルの決定係数が低い場合、そのモデルの回帰係数の信頼性は損なわれる
5. 2つの候補モデル A, B に対して AIC と BIC を用いてモデル選択を行う。いかなるサンプルサイズにおいても、AIC と BIC のモデル選択結果は一致する（ただし、モデル A, B それぞれの対数尤度および推定パラメータ数はサンプルサイズに依らず一定と仮定する）

小問 2

データのサンプリング方式について、次の選択肢から正しいものをすべて選択してください。

1. 有限母集団から母集団の半分の数だけ非復元ランダムサンプリングを行った場合、母集団の大きさに依らず標本平均の分散は常に等しくなる
2. 有限母集団から非復元ランダムサンプリングを行った標本平均の分散は母集団を無限母集団とみなしても同じである
3. 有限母集団から復元抽出した標本平均の分散は母集団を無限母集団とみなした場合と同じである

小問 3

あるクライアントからサーバ改修の依頼を受けました。クライアントにヒアリングしたところ、「既存サーバの平均的な応答速度については問題ないが、応答時間の分散がより小さくなるようにサーバ改修をお願いしたい」という要望が見えてきました。そこで既存サーバ A に対して、新サーバ B を開発しました。それぞれのサーバの応答速度について 16 個ずつデータを取得したところ、標本平均からの偏差平方和はサーバ A が 180、サーバ B が 90 でした。ただし、サーバ A・B の応答時間はそれぞれ独立かつ正規分布に従うものとします。

[3-1] サーバ B における応答速度の母分散の 95% 信頼区間として最も適切なものを次から選択してください。

1. 信頼下限: 4.00, 信頼上限: 16.00
2. 信頼下限: 2.78, 信頼上限: 16.78

3. 信頼下限: 1.48, 信頼上限: 28.96

4. 信頼下限: 3.27, 信頼上限: 14.38

5. 信頼下限: 5.68, 信頼上限: 18.34

[3-2] サーバ A よりもサーバ B のほうが応答時間の分散が小さいという主張を検証するため、有意水準 5% で片側検定を実施することになりました。サーバ A の応答時間の分散を σ_A^2 、サーバ B の応答時間の分散を σ_B^2 とします。この場合

(a) 適切な検定手法は何か

(b) 「サーバ A とサーバ B の応答時間の分散に差はない」という帰無仮説は棄却できるか

(c) 応答時間の分散について、 $\sigma_A^2 > \sigma_B^2$ ではなく、 $\sigma_B^2 = t$ (t は特定の値) であると主張したい場合に適切な検定手法は何か

に対する解答の組として適切なものを選択してください。

1. (a): t 検定, (b): 棄却できる, (c): カイ二乗検定
2. (a): カイ二乗検定, (b): 棄却できる, (c): t 検定
3. (a): F 検定, (b): 棄却できる, (c): カイ二乗検定
4. (a): t 検定, (b): 棄却できない, (c): カイ二乗検定
5. (a): カイ二乗検定, (b): 棄却できない, (c): t 検定
6. (a): F 検定, (b): 棄却できない, (c): カイ二乗検定

小問 4

あるサイトのデザイン変更施策を担当することになり、新デザインと旧デザインを比較する官能評価実験を実施しました。実験は 6 名を対象とし、ランダムに 3 名ずつの 2 グループに分けました。片方のグループには新デザインを、もう片方のグループには旧デザインを提示し、デザインの良さについて 10 点満点で採点してもらいました。下記がその結果になります。

	1 人目	2 人目	3 人目
新デザイン	4	9	8
旧デザイン	7	3	2

[4-1] この結果に対して、「新デザインは旧デザインより採点結果が高い」という仮説を検証します。ウィルコクソンの順位和検定を行った場合の片側 p 値（境界値を含める）として適切なものを選択してください。

1. 0.01
2. 0.05
3. 0.1
4. 0.2

5. 0.5

[4-2] この実験に対して、サンプルサイズが少なすぎるのではないかという指摘が入りました。有意水準（境界値を含める）を 0.05 とした場合、この実験に必要な最小サンプルサイズは何人か選択してください。

1. 5
2. 6
3. 8
4. 10
5. 12
6. 32

小問 5

飲食店の売上とネット掲載有無の関連性について調査を実施することになりました。まず、直近半年で売上が減少した店舗のうち 12 店舗を非復元サンプリングしました。更に比較のために、サンプリングしたそれぞれの店舗と規模や業態が似ていて、且つ売上が減少していない店舗を各々 3 店舗ずつ非復元サンプリングしました。店舗のネット掲載有無が売上増減の決め手になっているのではないかという仮説をもとに、下記のクロス集計を実施しました。

	売上減少有り	売上減少無し	合計
ネット掲載有り	8	30	38
ネット掲載無し	4	6	10
合計	12	36	48

この集計から得られる解釈として正しいものをすべて選択してください。ただし、ここで言及する売上減少率とは (売上が減少した店舗数) / (売上が減少した店舗数 + 売上が減少していない店舗数) を指しています。また、サンプリングに問題はないこととします。

1. ネット掲載無し店舗の売上減少率は 40%、ネット掲載有りでの売上減少率は 21% である。これはネット掲載有無による売上減少率の妥当な推定値といえる
2. ネット掲載無し店舗の売上減少率は 40%、ネット掲載有りでの売上減少率は 21% である。これはネット掲載有無によって売上減少率が 19% 削減できたということを示している
3. この分割表に対し、有意水準 5% でカイ二乗検定を行ったところ有意差はみられなかった
4. この分割表に対し、オッズ比を求めたところ 2.5 であった。これはネット掲載有無による売上減少率の相対比はおおよそ 2.5 であると推定できる