

2026年度

大学院文学研究科博士課程前期2年の課程入学試験

(冬期・一般選抜) 問題

専門科目 計算人文社会学 専攻分野

試験開始の合図があるまで、この問題冊子を開いてはいけない。

受験記号番号	
--------	--

成	
績	

2026 年度

大学院文学研究科博士課程前期 2 年の課程入学試験

(冬期・一般選抜) 問題

専門科目 (計算人文社会学 専攻分野)

注意) 解答の順序は自由であるが、どの問題の解答であるかが分かるように、問題番号を間違いなく記入すること。

問題 1. 観測データ $(x_1, y_1), \dots, (x_n, y_n)$ に対して、以下の線形回帰モデルを仮定する。

$$y_i = \alpha + \beta x_i + u_i, i = 1, \dots, n.$$

ここで、 u_i は非確率的な残差である。残差 u_i の二乗和を最小化する α, β の推定値を求めよ。計算過程も示すこと。

(次頁に続く)

問題2. 図1の 2×2 のクロス表に関して、以下の問いに答えなさい。

	1	2	計
1	n_{11}	n_{12}	$n_{1.}$
2	n_{21}	n_{22}	$n_{2.}$
計	$n_{.1}$	$n_{.2}$	N

図1 2×2 のクロス表

- (1) 変数1と変数2が完全に関係している状態を表中の記号(例: n_{11})で表現せよ。
- (2) 変数1と変数2が独立している状態を表中の記号(例: n_{11})で表現せよ。
- (3) 前問(1)のような関係にあてはまる変数1と変数2の現実社会の例を示しなさい。

問題3. 次の語句について、4つから2つを選択し、1語句につき100字程度で簡潔に説明せよ。

- ①Word2Vec ②媒介中心性 ③事前学習 ④ハイパーパラメータ

問題4. 以下の英文を読み、問いに答えなさい。

- (1) 下線部 (a) について、一般的に機械学習の分野において、データを”training set”と”testing set”に分割するのはなぜか。
- (2) 下線部 (b) を日本語に訳しなさい。
- (3) 下線部 (c) で述べられているコンピュータ科学者と社会学者の違いは、社会学者の側には何をもたらすかについて、筆者の主張をまとめなさい。

著作権の都合上、この部分をご覧いただけません。

著作権の都合上、この部分をご覧いただけません。

(出典) DiMaggio, Paul. 2015. "Adapting Computational Text Analysis to Social Science (and vice versa)." *Big Data and Society* 2(2): 1-5. (出題範囲は pp. 1-2 に基づく)
