

# 教育・広報に役立つデータサイエンス統計研修

—AI・データサイエンス学科構築研修第二弾—

## 研修の日程

### 第1日目 2020年2月20日(木)

時間	概要	詳細内容
12:30~13:00	受付	
13:00~13:05 (5分)	1日目の開会挨拶 研修スケジュールの案内	
13:05~14:20 (75分)	<ul style="list-style-type: none"> <li>研修の概要</li> <li>データサイエンス, 統計を学ぶ価値</li> </ul> 第1章「データを分布で捉える」 <b>【Keyword】</b> 記述統計, 代表値, 分散, 標準偏差, ヒストグラム, 正規分布 第2章「相関分析」 <b>【Keyword】</b> 散布図, 相関係数, 相関と因果	<p>「データを分布で捉える」ことの重要性を理解し、Excel と R によって分布を表現する各種技法を習得する。</p> <p>実際に専門学校でデータサイエンスの授業を担当することを考慮し、関数の入力だけでなく、その意味や理論についても演習を通じて理解できるよう心掛ける。</p> <p>また、相関及び因果の関係に注目することを通して、数値やグラフ以外の角度からもデータを分析できるよう視点を広げていく。</p>
14:20~14:30 (10分)	休憩	
14:30~15:30 (60分)	第3章「分析」から「予測」へ <b>【Keyword】</b> 単回帰分析, 重回帰分析, 質的変数, 量的変数, 数量化I類, 多重共線性	<p>データ活用の視点を「分析」から「予測」へと広げていく。Excel を用いた回帰分析の技法を習得し、務上の注意点等についても併せて解説する。</p>
15:30~15:40 (10分)	休憩	
15:40~16:55 (75分)	第4章「仮説検定」 <b>【Keyword】</b> 標本と母集団, 推測統計, 二項検定, 有意差, p値, $\chi^2$ 検定	<p>推測統計の山場でもあり魅力でもある「仮説検定」を扱う。簡単な実験及び PC を活用したシミュレーションを通して、「検定」の概念について正確に理解する。</p> <p>その後、実務応用として <math>\chi^2</math> 検定を Excel と R で行う。また、第1限、第2限で扱った分析についても検定の視点から再評価する。</p>
16:55~17:10 (15分)	休憩	
17:10~17:55 (45分)	第5章「多様なデータサイエンス」 <b>【Keyword】</b> テキストマイニング, 出版バイアス, サンプルバイアス	<p>数値以外の視点からデータを分析する手法としてテキストマイニングを取り上げ、アンケートの自由記述の分析を行う。</p> <p>また、「データの分析」だけでなく「データの収集」時点からデータサイエンスは始まっていることを説明し、収集時に陥りやすい各種のバイアスについて解説する。</p>
17:55~18:00 (5分)	事務連絡	
	(会場移動)	
18:30~20:30	情報交換会 (多様な学校の方と情報共有できる貴重な時間です)	

第2日目 2020年2月21日(金)

時間	概要	詳細内容
9:30~9:35	2日目の開会挨拶	
9:35~11:05 (90分)	第6章「種々の検定」 【Keyword】 t検定, 分散分析, 多重比較	$\chi^2$ 検定以外の様々な検定を取り扱う。1つ1つの検定のメカニズムについては深掘せず、種々の検定手法について概観できるようにする。実際の業務データを用いた検定の活用方法についても説明する。
11:05~11:15 (10分)	休憩	
11:15~12:25 (70分)	第7章「最新の統計事情」 【Keyword】 頻度論とベイズ論, 機械学習, SVM  第8章「カリキュラムの策定に向けて」 【Keyword】 高校情報科・数学科のデータサイエンスカリキュラム	今までの学習内容の振り返りを行った上で、データサイエンス学科の構築に主眼を置いた周辺事情を説明する。ベイズ統計学や機械学習について扱う予定である。 また、高等学校の学習指導要領改訂について触れ、今後専門学校の入学者が統計について、どのようなカリキュラムで、どのような内容を学んで入学するかを現場教員の立場から解説し、カリキュラム策定に活用できる形を取りたい。
12:25~12:30	閉会挨拶 アンケート	