

【一元配置分散分析を用いたアンケート分析】

【分析の目的】

本レポートの分析対象は、素材メーカーの研究員向けに実施している「MI 解析 OJT」という教育プログラムである。これは MI 解析手法をレクチャーする 3 か月間の OJT プログラムである。講義の各回でアンケートを取っており、各内容の理解度を 5 段階で集計している。今回のレポートでは、このアンケート結果を使用し、回答者によって理解度にバラつきが生じるかを検定する。

教育プログラムの運営として目指すのは、全ての部署の研究員が MI を使いこなせる様にする事である。その為には、受講者のドメイン(所属部署や研究分野)によって理解度に差が出ない様な講義内容にしなければならない。例えば、OJT 受講者にはポリマー加工をしている研究員や製造工程を検討している研究などがあり、背景に持つ学問は様々なので、どの様な背景を持つ受講者にも当てはまる解析技術を伝える必要があるし、表現や伝え方も考慮しなければならない。今回の分散分析によって有意差が出てしまうという事は、受講者によって理解のしやすさに偏りが出ている事を示唆する。実際に統計的仮説検定を用いてこれを確かめる。

因みに理解度を左右する要因には、上記の「分かり易い講義の実施」以外にも複数存在する。例えば受講者毎の前提知識、主観的な採点基準、などが挙げられる。本来はこれらも考慮する為に、多元配置分散分析を行うと良いのだが、今回はポートフォリオ用のレポートなので、自身の業務イメージを伝える事を優先して、明快な1元配置分散分析を行う。

【分析対象 —MI 解析 OJT— について】

MI 解析 OJT は「CRISP-DM」という普遍的解析サイクルに沿って講義を行うプログラムであり、「ビジネス理解」→「データ理解」→「データ準備」→「モデル構築」→「モデル評価」→「共有・展開」というステップの順で講義は進む。

また、各講義の終了後に受講者にはアンケート回答を依頼しており、各講義内容についての「理解度」を 5 段階で回答頂いている。

図 1: Microsoft Forms を使って受講後アンケートを実施している図



【一元配置分散分析結果】

アンケート回答者 7 名、1 人あたりの回答数42項、合計 294 行のデータを使って1元配置分散分析を行った。

アンケート結果は、全体の平均値＝4.505669、全体の偏差平方和＝61.24 となった。また、回答者別の偏差平方和＝37.92952 となったので、測定誤差の偏差平方和＝23.31651 となった。従って、分散値(F 値)＝7.179 となり、優位水準5%で有意といえる。ただし、優位水準1%だとギリギリ有意といえない。この結果は、教育プログラム運営をしている私としても、妥当に感じる。

図2: 分散分析表

	平方和	自由度	分散	分散比(F)	F(0.05)	F(0.01)
全体	61.2426	35	1.74978			
回答者	36.6011	6	6.10019	7.17919305	3.788	7.1799
誤差	24.6414	29	0.84970			

【分析結果の考察・感想】

この検定結果は、妥当に感じた。アンケート結果の平均値を回答者別に示すと、以下の表のとおりになる。

表2: 受講者のアンケート結果の平均値

参加者	平均値
K さん	4.97619
T さん	4.944444
M さん	4.095238
TN さん	4.333333
S さん	4.761905

O さん	4.380952
SZ さん	4.047619

直感的に分かり易い平均値比較をしてみても、受講者によってバラツキがありそうだ、というのは見て取れる。幸いなことは、全ての受講者の理解度が 5 段階中 4 以上であった事だ。理解度にバラつきはあれど、概ねは全体にとって分かり易い講義が出来ていると見て良いだろう。

【注釈】

このレポートで用いた分散分析は要因が有意かどうかを検定するものであり、各要因の効果量測定までは行わない為、「検定に費やした時間と労力の割には大した発見は得られなかった。」という結論で終わってしまう。よって分散分析を行う時は、今回の様に1元配置で行うのではなく、多元配置分散分析を行う事が望ましいと考える。今後は「前提知識」や「専攻科目」や「キャリア年数」などを含めると、交互作用も考慮した検定を行うことが出来るので、より得られる発見が多いただろう。