

路面の体感評価試験における被験者数について

富山和也¹・川村 彰²・石田 樹³・高橋 清⁴

¹ 学生員 修(工) 北見工業大学大学院工学研究科博士後期課程 (〒090-8507 北海道北見市公園町 165 番地)
E-mail : tomiyama@vortex.civil.kitami-it.ac.jp

² 正会員 博(工) 北見工業大学教授 工学部社会環境工学科 (〒090-8507 北海道北見市公園町 165 番地)

³ 正会員 独) 土木研究所 寒地土木研究所 (〒062-8602 北海道札幌市豊平区平岸 1 条 3 丁目)

⁴ 正会員 工博 北見工業大学准教授 工学部社会環境工学科 (〒090-8507 北海道北見市公園町 165 番地)

本研究の目的は、路面平坦性に対する利用者許容度の体感評価において、統計的に必要な被験者数を検定力分析により算定する事である。被験者数は、統計的有意水準、母集団効果量、検出力の関数となるが、効果量については、個々の研究に委ねられる。そこで本研究では、ドライビングシミュレータによる体感評価試験から、平坦性評価における効果量を算出した。その結果、中から大の効果量は、車両上下加速度実効値 0.35m/s^2 以上の低減が見込める場合であり、この場合 20 名以上の被験者数でよい事がわかった。一方、加速度実効値の差が 0.3m/s^2 以下の小さな効果量では、400 名以上の被験者を要する場合がある。必要人数に満たない試験では、統計的信頼区間を求め、路面評価における体感試験結果の一般化を避けるべきである。

Key Words : pavement roughness, sample size, statistical power analysis, driving simulator, human factor

1. はじめに

舗装の設計や施工において性能規定が導入され、平坦性やわだち掘れといった路面性状に対しても要求性能の保証と確保の重要性がますます高まっている。平坦性やわだち掘れは、車両挙動に影響し乗り心地を低下させるなど、道路利用者の快適性や安全性に直結するため、利用者意識に基づく損傷度合いの評価が必要である。そのため、従来、利用者の体感に基づく乗り心地や安心感の評価試験が実施され¹⁾、とりわけ、今日のバーチャル・リアリティ技術の進化に伴い、ドライビングシミュレータ(以下、「DS」とする)を用い、安全かつ迅速に体感評価試験(以下「体感試験」とする)を実施する方法が注目されている²⁾。DSを用いた体感試験は、実道実車試験に比べて、被験者の安全性や試験条件の設定が容易な事から、今後もその需要は高まるものと考えられる。

このような背景から、筆者らの研究室では人と路面をつなぐインターフェイスとして、路面評価型の DS (KITDS: Kitami Institute of Technology Driving Simulator)を開発している³⁾。KITDSに固有の特徴として、路面プロファイルや路面損傷に起因する車両振動を、前方映像とともに再現することが可能である。現在、路面損傷に起因する車両振動の再現性に関するキャリブレーション作業が終了し、これまでに行われた体感試験結果から、路面評価への有効性が示されている^{1),4)}。

体感試験の主目的は、限られた被験者による路面の体

感評価値の一般化である。換言すれば、標本統計量の値をもとに、できるだけ正確な母数を推測する、という統計的推測の課題と一致する⁵⁾。路面損傷に起因する振動乗り心地の解析においても、しばしば統計的推測、とりわけ統計的検定が用いられる^{1),6)}。この場合、測定結果の標本統計量は、常にサンプルサイズ(被験者数)に依存し⁷⁾、被験者数が増えるに従い実験コストも増加する。そのため、実験計画段階で、分析目的や方法などの与条件に応じた方法で、最低限必要な被験者数を決定することが効率的かつ重要である^{8),9)}。しかしながら、過去にも路面の体感試験において被験者数が問題となる事はあったが、明確な根拠は不明である。本研究の目的は、利用者にとって路面平坦性に起因する車両振動が許容できるか否かという利用者の許容度(以下、「路面許容度」とする)に関する体感試験実施にあたり、母数を推測する上で重要な要因である被験者数について、統計的推測に基づき理論的に明らかにする事である。

統計的推測とは、推測統計や統計的推論とも呼ばれ、確率モデルに基づき、標本から母数に関する推測を行う方法であり、測定値の特徴を客観的に記述するための記述統計とは区別される。また、統計的推測は、統計的推定(以下、「推定」とする)と統計的検定(以下、「検定」とする)に大別される^{5),8),10)}。

被験者数を決める目的は、(a)推定において推定精度を確保したい場合と、(b)検定によりある仮説を統計学的に検定したい場合がある¹¹⁾。(a)推定の場合、一般に被験

者数が多くなれば、区間推定における信頼区間の幅が狭まり、より精度の高い推測が可能となる。一方、(b)検定の場合、極端に多い被験者数は、多くの被験者が改善効果を実感できないような意味のない差までも検出してしまい、その有意性ゆえに結果を過大評価しかねない¹²⁾。逆に、極端に少ない被験者数では、実際には改善効果があるものでも、検出力が低く、重要な結果を見落とす可能性がある。ここで、検出力とは帰無仮説が正しくないときに、それを正しく棄却できる確率である⁷⁾。

統計的推測に基づき被験者数を定める方法は、区間推定の場合、信頼区間の幅をどの程度に抑えたいかを考え、信頼区間を求めるための公式を利用して、必要な被験者数を定めることができる。ただし、信頼区間の幅は標本統計量の値に依存するため、予備調査や類似研究により予想する必要がある⁸⁾。一方、検定の場合、被験者数は、検出力、帰無仮説からの隔たりの程度である母集団における効果量（以下、「効果量」とする）、有意水準の関数となる⁷⁾。検出力や有意水準については、統計的に妥当な値が存在するが、効果量は研究対象や目的に応じて適当な値を定める必要がある¹¹⁾。また、実験段階では効果量は未知であるため、予備調査や既往研究などから推測する必要がある。このように、統計的に被験者数を決定する方法については一般化されているが、必要人数を決める上で重要な要素である、信頼区間の幅や検定における効果量については根拠が不明である。

そこで本研究では、2006年から2008年の間に行ったKITDSによるアンケート形式の体感試験を基に効果量を推測し、路面許容度の評価に必要な被験者数について検討する。とりわけ、被験者意識に基づく路面改善度判定を鑑み、異なる2水準の車両振動の差に着目した検討を行った。なお、区間推定と検定とは表裏一体の関係にあり¹⁰⁾、同様の研究において検定が多用されている事、信頼区間を求める推定に比べ検定の方が被験者数の影響が問題となる事を考慮し、検出力を求めるための検定力分析に基づき⁷⁾被験者数の算定を試みる。また、大学における体感試験では、被験者としてしばし学生に協力を依頼するが、学生被験者の評価を一般化できるかどうかについての検討も行う。統計手法の計算処理については、統計解析向けの言語・環境であるR¹³⁾、¹⁴⁾および検定力分析ソフトウェアのG*Power¹⁵⁾、¹⁶⁾を用いた。

2. DSによる体感評価試験

体感試験を実施するにあたり、実道における試験では、被験者の安全性や必要とされる路面条件の確保の観点から多くの困難をとまう。一方、DSを用いた体感試験では、安全性の確保や実験条件の容易な設定、同一条件で繰り返し試験が行えるなど多くの利点を有する事から、



図-1 KITDSの外観

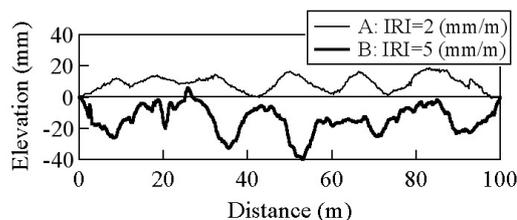


図-2 実験に用いた路面プロファイル
(EVEN 試験結果内 A: SITE=9, L=110-210m /
B: SITE=4, L=40-140m)

近年高い注目を集めている²⁾。KITDSは、安全運転教育や景観評価といった従来の機能に加え、実路面プロファイルや実車両の挙動を前方映像とともに再現可能な路面評価型のDSである(図-1)。KITDSにより、従来では困難とされていた路面損傷に起因する車両の振動乗り心地の評価を、安全かつ効率的に実施する事が可能となった。以下に、本研究で対象とするKITDSによる体感試験の概要を、主要な項目毎に示す。

(1) 被験者

本研究では、KITDSによる体感試験において、学生を除く北見市在住の20代から60代までの一般被験者男女28名による一般群及び、北見工業大学に在学する男子学生28名による学生群の2サンプル群を対象とした。なお、予備試験は実施せず、全ての被験者は、本試験で使用した走行シナリオの初体験者である。ここで、一般群と学生群において、被験者の属性以外の試験条件は全て等しくし、また、同一試験内において、被験者毎の室温や照明等の試験環境は極力等しくした。

(2) 路面プロファイルに起因する振動条件

平坦性に起因する振動条件は、車両のバネ上上下下加速度の実効値（以下、「加速度」とする）を基準に、IRI (International Roughness Index) 2水準 (IRI=2, 5mm/m)、速度2水準 (60, 80km/h) とした。ここで、IRI=2mm/mは新設舗装に相当し、IRI=5mm/mは供用後の舗装に相当する路面状況である。当該プロファイルは1998年に行われた路面平坦性に関する第2回PIARC国際共通試験（通称「EVEN試験」）の測定データを用いた¹⁵⁾。実験に使用

表-1 体感評価試験における振動条件

Case	IRI (mm/m) []内は図-2 に対応	走行速度 (km/h)	上下加速度実 効値(m/s ²)
1	2 [A]	60	0.34
2	2 [A]	80	0.48
3	5 [B]	60	0.86
4	5 [B]	80	1.13

したプロファイルを図-2 に示す。振動区間延長は、各プロファイル 100m を 3 度繰り返した 300m であり、前後に 10 秒の助走区間を設けた。以上の条件を車両挙動シミュレーションソフトウェアの CarSim[®]により計算し¹⁶⁾、KITDS の体験走行機能により振動及び前方映像を再現、体感試験を実施した。本機能は、実車両や理論的な車両モデルから得られた車両挙動を、前方映像と共に、六軸モーションベースシステムにより再現し、運転操作を伴わずに走行感覚の体験を可能とする。振動条件を表-1 に示す。なお、振動の提示順序は、振動の残留効果を低減するため、被験者毎に無作為に振り当てた。

(3) 評価尺度

本研究では、アンケートによる路面許容度の評価を実施し、各群共通の尺度として「許容度（許容できる・許容できない／どちらともいえない）」を用いた。本尺度は、従来から乗り心地の研究に用いられており¹⁷⁾、許容の可否を問題にしていることから、路面管理に反映可能な尺度であると考えられる。また、鈴木ら¹⁸⁾は鉄道の乗り心地評価において、振動の程度を「かなり」や「やや」といった副詞で表すよりも、日常の体験に即した「諾否」で問う方が、より絶対評価に近い回答が得られるとしている。そのため、本研究では、路面平たん性の「許容度」に着目した。

3. 平たん性の許容度評価における

効果量及び被験者数

本章では、平たん性の許容度評価において、車両振動に対する被験者の乗車感覚に着目し (1) 学生被験者の一般性、(2) 加速度差に基づく効果量及び体感試験に必要な被験者数について検討する。

(1) 学生被験者の一般性について

体感試験において、被験者に対しては振動という一種のストレスを与え、試験目的によってはプライバシー情報を記録するため、被験者に対する対応や、情報の取り扱いなど倫理的問題には最大限注意を払う必要がある。実験計画上、被験者募集は最も労力を要するところである。大学において体感試験を実施する際に、属性の異なる一般被験者を集める場合、学外のため、交通手段の確保や拘束時間などの問題や制約が生ずる。一方、学生被験者の場合、移動や時間の制約などについて、一般被験

表-2 体感評価試験における被験者の運転頻度

運転頻度	学生群 (人 [%])	一般群 (人 [%])
ほぼ毎日	9 [32]	22 [79]
1 週間に数回	3 [11]	2 [7]
1 ヶ月に数回	3 [11]	1 [4]
1 年に数回	7 [25]	1 [4]
ほとんどしない	6 [21]	2 [7]
合計	28 [100]	28 [100]

表-3 一般被験者の性別及び年代 (人)

年齢層 性別	20	30	40	50	60	70	合計
男性	2	1	3	4	4	0	14
女性	2	3	2	2	4	1	14
合計	4	4	5	6	8	1	28

者に比べ柔軟な対応が可能である。そのため、筆者らは、実験計画策定の初期段階における予備試験においては、学生を被験者とした試験を行ってきた。しかし、限られた属性の被験者群から得られた結果が代表性をもち一般化できるのか、といった問題がある。そこで、一般群と学生群による体感試験を比較することにより、路面許容度評価における学生被験者の一般性について検討した。ここで、学生被験者とは、学部及び大学院に所属する学生とし、一般被験者とは学生被験者を除いた者である。なお、本実験における被験者の概要は前章で述べた通りである。

a) 被験者の運転頻度

各群被験者の運転頻度を表-2 に示す。表-2 より、被験者のうち、学生群では 4 割強の被験者が週に数回運転しており、およそ半数は年に数回程度しか運転をしない。一方、一般群では 8 割以上の被験者が週に数回は運転しており、年に数回程度の運転頻度は全体の 1 割程度である。なお、被験者は全員運転免許を保有している。また、一般群の年齢層及び性別を表-3 に示す。

b) 学生被験者と一般被験者の許容度比較

許容度の比較条件は表-1 に示す 4 条件である。比較方法として、予備試験における学生の協力を鑑み、各振動条件に対し無作為に各群 7 名ずつを割当、対応のない 2 要因以上の比率の差の検定法である逆正弦変換法により許容率の差の検定を 5%水準で行った。逆正弦変換法とは、 n 個の検討対象とする比率 P に式 (1) の変換を施し、これを対象データの平均値 X とみなす事により分散分析を応用し、各要因の効果の検定を行う手法である¹⁰⁾。

$$X_i = \sin^{-1} \sqrt{P} \quad (i = 1, 2, 3, \dots, n) \quad (1)$$

また、比率が $P=1$ 及び $P=0$ の場合には、以下の補正比率を用いることが望ましいとされる¹⁰⁾。

$$P = 1 \Rightarrow P = \frac{n - 0.25}{n} \quad (2)$$

表-4 各振動条件に対する許容度及び許容率

Case	1: IRI=2mm/m, V=60km/h		2: IRI=2mm/m, V=80km/h		3: IRI=5mm/m, V=60km/h		4: IRI=5mm/m, V=80km/h		合計
	学生	一般	学生	一般	学生	一般	学生	一般	
許容	7	6	4	7	3	5	2	2	36
その他	0	1	3	0	4	2	5	5	20
合計	7	7	7	7	7	7	7	7	56
許容率	1.000	0.875	0.571	1.000	0.429	0.714	0.286	0.286	0.643
補正比率	0.964	0.857	0.571	0.964	0.429	0.714	0.286	0.286	0.643

*許容度の単位：人

表-5 逆正弦変換法による分散分析表

変動因	平方和	自由度	χ^2	p 値
振動条件の主効果	1940.993	3	16.549	0.000
属性の主効果	157.361	1	1.342	0.247
振動-属性の交互作用	497.682	3	4.243	0.236
群内分散	117.286			

表-6 各振動条件間における上下加速度実効値の差 (m/s²)

	Case 1	Case 2	Case 3	Case 4
Case 1	-	0.14	0.52	0.79
Case 2	-	-	0.39	0.65
Case 3	-	-	-	0.27

表-7 各振動条件間における許容度の変化と比較振動の許容率

Case	(対象振動, 比較振動)	(1,0)* (人)	(0,1)* (人)	変化合計 (人)	変化人数が全体に占める割合 (%)	比較振動の許容率 (%)
比較 1	(Case1, Case2)	5	2	7	12.5	28.6
比較 2	(Case1, Case3)	27	1	28	50.0	3.5
比較 3	(Case1, Case4)	30	0	30	53.5	0
比較 4	(Case2, Case3)	25	2	27	48.2	7.4
比較 5	(Case2, Case4)	27	0	27	48.2	0
比較 6	(Case3, Case4)	11	7	18	32.1	38.9

*1: 許容, 0: その他

$$P = 0 \Rightarrow P = \frac{0.25}{n} \quad (3)$$

なお、評価尺度において、「許容できる」のみに着目し、「許容できない」及び「どちらともいえない」は「その他」として扱った。許容率の集計表を表-4に示す。

表-4のデータについて、逆正弦変換法により得られた分散分析表を表-5に示す。表-5より振動条件の主効果は5%水準で有意である。一方、本検討で対象とする属性の主効果は、5%水準において有意な差はみられなかった。また、この時の検出力をCohenの方法⁷⁾によりRを用いて算出したところ81%であった。また、交互作用も、5%水準で有意ではなく、この時の検出力は99%である。このことから、許容度の評価においては、学生群と一般群の統計的な差は大きくないことがわかる。よって、次節以降では、学生群と一般群を合わせた被験者群を対象とする。ただし、乗り心地評価においては、被験者の体力的な要因が評価に影響を及ぼすケースも少なくないと考えられる¹⁸⁾。このような場合、限られた属性の被験者による体感試験は、予備試験に限定して実施すべきである。

(2) 振動条件間における路面許容度の検討

路面評価の主要な問題として、平坦人性水準の差や速度水準の差によって、許容度がどの程度異なるかが問わ

れる。これにより、例えば、路面の補修による改善効果の事前確認が可能となる。そこで、本節では、各振動条件間の加速度差をもとに、効果量を予測し、ある加速度差を検出するために必要な被験者数について検討する。なお、これ以降、標本から得られた効果量を母集団における効果量と区別するために、「標本効果量」と記し、標本効果量を効果量の推定値とする。

a) 振動条件間における許容率

それぞれの振動条件間における加速度の差を表-6に示す。表-6より、加速度差は0.14m/s²から0.79m/s²の間に、概ね0.1m/s²刻みで得られた。また、本検討では、各振動条件間での効果量に着目し、それぞれの比較条件を独立事象として扱う。各条件間の比較は対応のある2群比率の差の検定となり、McNemar検定や二項検定の適用が考えられるが¹⁰⁾、本研究では、正確率検定である後者を両側検定で用いた。なお、全振動条件間で許容率の差の検討を行う場合には、検定の多重性を考慮した手法を適用する必要がある。ここで、対応のある2群比率の差の検定において許容度の変化とは、許容を1、その他を0としたとき、ある条件(A,B)について、(1,0)から(0,1)、もしくはその逆となった度数である。表-7に各振動条件間の許容度の変化度数、及び変化した被験者が全体に占める割合、比較振動の許容率を示す。なお、便宜上、表-7

表-8 各振動条件間における標本効果量

比較群	(対象振動, 比較振動)	加速度実効値の差 (m/s^2)	標本効果量
比較1	(Case1, Case2)	0.14	0.21
比較2	(Case1, Case3)	0.52	0.46
比較4	(Case2, Case3)	0.39	0.43
比較6	(Case3, Case4)	0.27	0.11

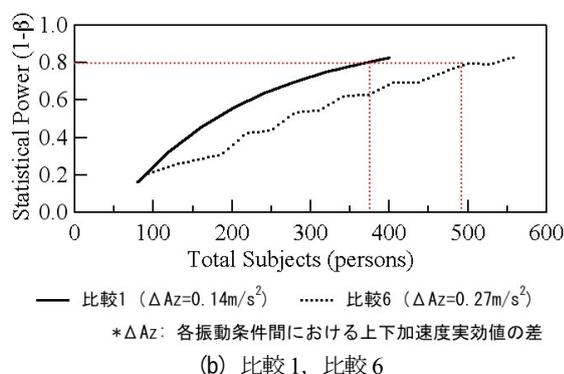
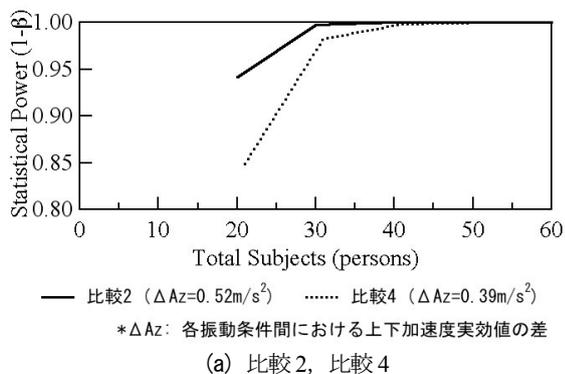


図-3 各振動条件間における検出力曲線

に示す通り、振動条件の一方を対象振動、もう一方を比較振動とした。

表-7より、比較3と比較5では比較振動の許容率が0%となっている。これらの比較における加速度差は、表-6よりそれぞれ $0.79m/s^2$ と $0.65m/s^2$ であり、振動条件間にこの程度の差がある場合に、必要な被験者数が10人程度でも検出力は100%となる。換言すれば、加速度を $0.65m/s^2$ 以上低減できるような改善策を施せば、かなり大きな改善効果が得られる事がわかる。逆に、この程度の差がある場合、体感試験から路面の許容度に関して得られる情報は少ないといえる。そこで、次節以降の解析において、比較3,5は除外した。

b) 効果量の推定及び被験者数の検討

2群比率の差の検定を二項検定により行う場合の帰無仮説 H_0 は母集団における比較振動の許容率 P が対象振動のそれと等しい: $P=0.5$ であり、対立仮説 H_1 は2群比率が等しくない: $P \neq 0.5$ (両側検定)である。ここで、効果量 g は次式により与えられる。

$$g = P - 0.5 \quad \text{or} \quad 0.5 - P \quad (4)$$

表-9 上下加速度の差に基づく路面評価に必要な被験者数の目安

加速度実効値の差 (m/s^2)	母集団効果量	効果の大きさ	被験者数の目安
0.7以上	0.45以上	大	10名以上
0.35~0.7	0.40程度	中	20名以上
0.35以下	0.30以下	小	400名以上

表-8に各振動条件間の標本効果量を示す。また、参考値として加速度差も合わせて示した。ここで、必要被験者数を検定力分析により決める方法は、Cohen⁷⁾に代表される検出力表や統計解析ソフトを利用する。検出力表では、表に与えられた値以外は補間による近似となるため、本研究では、Cohenの方法に基づき、検出力表の代わりにG*Powerを用い被験者数を計算した。Cohenの方法とは、有意水準、効果量、検出力を設定し、検出力表から被験者数を決定する手法である⁷⁾。なお、有意水準 $\alpha=0.5$ 、検出力 $1-\beta=0.8$ とする。これらの値は、統計的に概ね妥当とされる値である¹¹⁾。図-3に、必要被験者数を求めるための、各振動条件間の標本効果量における検出力曲線を示す。図-3(a)より、比較2(加速度差 $0.52m/s^2$)及び比較4(加速度差 $0.39m/s^2$)では、被験者数が20名程度であっても検出力80%以上を確保できる。一方、図-3(b)より、比較1(加速度差 $0.14m/s^2$)では被験者数約400名、比較6(加速度差 $0.27m/s^2$)では被験者数約500名を超えなければ検出力80%を確保できない。このことから、体感試験において $0.27m/s^2$ 以下の加速度差を評価するためには、約400人を超える被験者が必要であることがわかる。ただし、振動条件間で評価が変化した被験者の割合は、表-7より、比較1で12.5%、比較6で32.1%と、3割程度しかおらず、この程度の加速度差を基に路面を補修しても、利用者の体感に基づく平たん性の大きな改善効果は見込めないものと考えられる。

c) 考察

体感試験から得られた標本効果量を基に検出力曲線を求めた結果、振動条件間に $0.39m/s^2$ 程度の加速度差があれば、20名程度の被験者数であっても、高い検出力を確保できる事がわかった。一方、 $0.27m/s^2$ 程度の加速度差を検出する場合には、400名以上の被験者が必要である。数百名規模の被験者を要する体感試験を頻繁に実施するのは現実的でない。また、事前に算出した被験者数より少人数での試験から得られる情報が、必ずしもゼロであるとは限らない。そのため、所用の被験者数に満たない場合や、微小な加速度差を研究対象とする体感試験では、検定ではなく推定により信頼区間を求め¹⁹⁾、結果についての一般化や、管理水準値策定などの重大な結論をさけるべきである。上述の結果を基に、加速度差に基づく路面の体感試験の場合に要する被験者数及び効果量の目安を表-9に示す。表-9において、「効果の大きさ」とは、

加速度差をなくした場合に見込まれる平坦性の改善効果である。また、表中「加速度実効値の差」及び「母集団効果量」の基準については、以下に基づき設定した。

- ・効果の大きさ「大」では、加速度実効値の差について、比較5(加速度差 0.65m/s^2)を下限とし、安全側をとり 0.7m/s^2 とする。母集団効果量は、比較4(標本効果量0.43)が効果の大きさ「中」の上限とし、安全側をとり0.45以上とする。
- ・効果の大きさ「中」では、加速度実効値の差について、比較4(加速度差 0.39m/s^2)と比較6(加速度差 0.27m/s^2)の平均 0.325m/s^2 の安全側 0.35m/s^2 を下限とする。母集団効果量は、比較4(標本効果量0.43)から0.40程度とする。
- ・効果の大きさ「小」では、加速度実効値の差について、効果の大きさ「中」の下限値より、 0.35m/s^2 以下とする。母集団効果量は、比較1(標本効果量0.21)と比較4(標本効果量0.43)の平均値0.32から0.30以下とする。

4. まとめ

本研究では、従来、理論的な根拠に乏しい体感試験における被験者数について、統計的検出力に基づき算定する方法を示した。これまで、アンケートによる路面乗り心地評価は客観性に欠けるとされてきたが、本研究はアンケートによる路面評価の信頼性や妥当性を補うものである。本研究のまとめを以下に示す。

- ・路面許容度に関して、KITDSによる2つの体感試験の概要を述べた。DSを用いた体感試験は、安全かつ効率的であり今後もその需要は高まるものと思われる。
- ・学生群と一般群の路面許容度の違いについて、逆正弦変換法により検討した結果、評価の傾向に大きな差はない事がわかった。ただし、試験条件による、体力的な影響が想定される場合、限られた属性の被験者による体感試験は予備調査に限定すべきである。
- ・振動条件間の加速度の差に着目した路面許容度の体感試験における、効果量及び効果の大きさ、最低限必要な被験者数を算出した。その結果、中から大の効果が見込まれる場合、即ち加速度差が 0.39m/s^2 以上(比較4)ある場合には、20名程度の被験者であっても、80%以上の検出力を確保できる事がわかった。一方、加速度差 0.27m/s^2 以下(比較6)の小さな効果を80%の確率で検出するためには400名を超える結果となった。400名を超す規模の体感試験は現実的でないので、事前に見積もった被験者数が得られない場合、推定により信頼区間を求め、路面評価において試験結果の一般化を避けるべきである。

参考文献

- 1) 石田 樹, 岳本秀人, 川村 彰, 白川龍生: ドライビングシミュレータによる舗装路面の乗心地・安心感評価, 土木学会舗装工学論文集, pp.49-56, 2004.
- 2) 川村 彰: ドライビングシミュレータを用いた道路研究の可能性, 道路建設, No. 687, pp.62-63, 2005.
- 3) Kawamura, A., Shirakawa, T. and Maeda, T.: KIT Driving Simulator for Road Surface Evaluation, *Proceedings of the 5th Symposium on Pavement Surface Characteristics*, CD-ROM, 2004.
- 4) Ishida, T., Shirakawa, T., Kawamura, A. and Tako, J.: Using the KIT driving simulator to evaluate road surface roughness, *Proceedings of the TRB 86th Annual Meeting*, CD-ROM, 2007.
- 5) 南風原朝和: 心理統計学の基礎—統合的理解のために, 有斐閣, 2002.
- 6) Ishida, T., Kawamura, A. and Tomiyama, K.: Quantitative Evaluation of Ride Comfort Using a Driving Simulator, *Proceedings of the TRB 87th Annual Meeting*, CD-ROM, 2008.
- 7) Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Assoc Inc, 1988.
- 8) 芝 祐順, 南風原朝和: 行動科学における統計解析法, 東京大学出版会, 1990.
- 9) 永田 靖: サンプルサイズの決め方, 朝倉書店, 2003.
- 10) 森 敏昭, 吉田寿夫: 心理学のためのデータ解析テクニカルブック, 北大路書房, 1990.
- 11) 吉田寿夫: 心理学研究法の新しいかたち, 誠信書房, 2006.
- 12) 杉澤武俊: 教育心理学研究における統計的検定の検定力, 教育心理学研究, 第47巻, 第12号, 1999.
- 13) 山田剛史, 杉澤武俊, 村井潤一郎: Rによるやさしい統計学, オーム社, 2008.
- 14) Faul, F., Erdfelder, E., Lang, A. G. and Buchner, A.: G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences, *Behavior Research Methods*, Vol. 39, No. 2, pp. 175-191, 2007.
- 15) Kawamura, A., Takahashi, M. and Inoue, T.: Basic Analysis of Measurement Data from Japan in EVEN Project, *Transportation Research Record*, No.1764, pp.232-242, 2001.
- 16) Mechanical Simulation Corporation: *CarSim Reference Manual*, 2006.
- 17) Sayers, M. W. and Karamihas, S. M. : *The Little Book of Profiling, - Basic Information about Measuring and Interpreting Road Profiles*, The University of Michigan, 1998.
- 18) 鈴木浩明, 白戸宏明: 体感による乗り心地評価試験の実施方法, 鉄道総研報告, Vol.17, No.1, pp.7-10, 2003.
- 19) 舟喜 光一, 折笠 秀樹: 信頼性の統計学—信頼区間および統計ガイドライン, サイエンス社, 2001.

DETERMINING THE SAMPLE SIZE OF SUBJECTS ON A DRIVING EXPERIMENT TO EVALUATE PASSENGER ACCEPTANCE OF PAVEMENT ROUGHNESS

Kazuya TOMIYAMA, Akira KAWAMURA, Tateki ISHIDA and Kiyoshi TAKAHASHI

The purpose of this study is to find statistically-significant number of subjects in a driving experiment that investigates the level of passenger acceptance of pavement roughness. Statistical power analysis is used to determine an appropriate sample size as a function of a significant level, effect size (ES) and statistical power. In power analysis, the ES basically depends on the purpose of each individual study. Therefore, we estimate the ES to evaluate the acceptance ratio of roughness based on a subjective rating by a driving simulator. Consequently, over 20 subjects in the experiment allows one to compare the acceptable proportion of two vertical accelerations split by more than 0.35m/s^2 in the root-mean-square value that is medium or large ES. Conversely, the small ES offered by the acceleration difference of less than 0.3m/s^2 requires a large number of subjects such as 400. If the required number of subjects is not obtained, we recommend that the statistical confidence interval should be calculated and the result of the experiment should not be generalized for the road surface evaluation.