Messexperiment Gruppengeschwindigkeit
1. Projekt zu Modellierung und Simulation

Prof. Dr. Gerta Köster

Sommersemester 2017
Ausgabe und Experiment am Mittwoch, den 05.04.2017
Abgabe am Freitag, den 19.04.2017, 15:15

1 Hintergrund: Modellierung Personenströme, Wunschgeschwindigkeit, „free-flow velocity“

Personenstromsimulatoren werden heute z.B. bei der Gebäudeplanung eingesetzt, um eine möglichst schnelle Räumung bei Feueralarm zu erreichen.

Dabei ist das Laufverhalten von Menschen noch ungenügend untersucht. Die Modell stützen sich oft auf bestenfalls plausible Annahmen. Das Laufverhalten auf Treppen ist einer der weitgehend weißen Flecken.

Von Individuen ist bekannt, dass sie *bei freier Bahn in der Ebene* bevorzugt mit ihrer persönlichen Wunschgeschwindigkeit, der sogenannten „free-flow velocity“ laufen. Man geht davon aus, dass diese normalverteilt ist um einen Mittelwert μ mit Standardabweichung σ [5]. Die konkreten Werte für μ und σ sind szenarienabhägig [1].

Aber wie hängt die Fortbewegungsgeschwindigkeit auf einer Treppe von der Wunschgeschwindkigkeiten der Probanden in der Ebene ab? Besteht überhaupt ein Zusammenhang? Wenn ja welcher? Viel diskutierte, aber unbewiesene und sich widersprechende Hypothesen sind, dass

- die Geschwindigkeit auf der Treppe linear von der Wunschgeschwindigkeit abhänge - nur etwas langsamer,
- und dass wegen der Taktung des Gehens durch die Stufen, kein Zusammenhang mit mit der Wunschgeschwindigkeit bestehe.

Sie sollen einen Beitrag zur Klärung leisten, indem Sie ein Messexperiment planen, ausführen und auswerten.

2 Aufgabe: Experiment und Auswertung

2.1 Messexperiment

Das Messexperiment findet in der Übung auf dem Gelände der Hochschule in der Lothstraße 64 statt. Für diesen Termin herrscht **Anwesenheitspflicht!** Ihre Teilnahme fließt in die Bewertung ein.

Eine "Choreographie“ wird Ihnen zur Verfügung gestellt. Sie bekommen Verstärkung durch Ihre Professorin.

1. Das Testszenario ist im Lichthof des R-gebäudes.

2. Messen Sie jeweils Ihre individuelle Wunschgeschwindigkeit in der Ebene. Messen Sie dreimal pro Nase und bilden Sie jeweils den Mittelwert als Schätzwert. (Stoppuhren stehen zur Verfügung. Alternativ nehmen Sie Ihre Smartphones.)

3. Messen Sie die mittlere Geschwindigkeit der Probanden entlang der Teststrecke auf der Treppe. Machen Sie auch hier 3 Runden. Sie erhalten so eine Reihe von Messpunkten, die Sie in Tabellen eintragen. Die erste Tabelle soll *Personen Id, Runden Id, Zeit, Treppe oder Ebene* enhalten. Die zweite Tabelle enthält die Daten der Personen mit den Feldern: *Personen Id, Geschlecht, Körpergröße.*

4. **Halten Sie sich beim Notieren der Messwerte und Mittelwerte streng an das Datenformat der Daten von 2012, die Sie auf dem Repo finden!**

2.2 Statistische Untersuchung

Für Ihre statistische Untersuchung nehmen Sie Mathematica oder Python oder – falls Sie sich schon damit auskennen oder es kennenlernen wollen – R (frei verfügbar).

1. Überprüfen Sie, ob Ihre individuellen Wunschgeschwindigkeiten normalverteilt sind. Mathematica bietet Ihnen dazu eine Reihe von sehr leistungsfähingen Tools: `Histrogram[]`, `QuantilePlot[]`, `DistributionFitTest[]`. Machen Sie sich klar, was diese Tools bieten und setzen Sie sie ein. In [3] finden Sie eine gute Einführung in eine Reihe geeigneter Tests wie den Shapiro-Wilk-Test. Zu welchem Schluss kommen Sie?

2. Über lineare Regression sollen Hinweise gefunden werden, ob ein Zusammenhang zwischen Treppengeschwindigkeit und den Größen Wunschgeschwindigkeit, Körpergröße und Rundennummer besteht. Legen Sie dazu jeweils eine Regressionsgerade durch die Messpunkte. Beginnen Sie mit einfachen Modellen, in denen die nur die Abhängigkeit von nur einer Variable gleichzeitig betrachten.

(a) Sie untersuchen also, ob ein linearer Zusammenhang der Gestalt $z = \beta_0 + \beta_1 x$ zwischen Treppengeschwindigkeit und Wunschgeschwindigkeit besteht. Dann wiederholen Sie die Auswertung für Körpergröße, dann Rundennummer.

(b) Betrachten Sie nun ein Regressionsmodell in zwei Variablen, der Wunschgeschwindigkeit und der Körpergröße. Sie untersuchen also einen linearen Zusammenhang der Gestalt $z = \beta_0 + \beta_1 x + \beta_2 y$.

(c) Ziehen Sie zuletzt noch die Rundennummer hinzu. Betrachten Sie also ein Regressionsmodell in drei Variablen.

Hinweis: Die Modelle heißen linear, weil die Regressionskurven linear in $\beta_\mathbf{1}$ und in $\beta_\mathbf{2}$ bzw. $\beta_\mathbf{3}$ sind!!!! Auch für die lineare Regression bietet Mathematica leistungsfähige Tools: `LinearModelFit[]`. Die Prinzipien lesen Sie z.B. in [4] oder [3] oder [2] nach. Machen Sie sich klar, was Mathematica (bzw. Ihr Pythonprogramm) ausrechnet.

3. Führen Sie für jedes der Regressionsmodelle einen t-Test durch, um zu überprüfen, ob tatsächlich ein linearer Zusammenhang besteht. D.h. stellen Sie die Nullhypothese H_0 auf, dass *kein* solcher Zusammenhang bestehe, also $\beta_1 = 0$ bzw. $\beta_2 = 0$ usw.. Wenn Sie die Regression in Mathematica mit dem Befehl `LinearModelFit[]` durchgeführt haben, können Sie die Testergebnisse über die "Property" `ParameterTable` direkt abfragen. Verwerfen Sie die Nullhypothese aufgrund des Ergebnisses oder behalten Sie sie bei? Erklären Sie, was die Einträge in der Tabelle bedeuten! Argumentieren Sie!

4. **Zusatzpunkte:** Untersuchen Sie die Konditionierung des Problems und beurteilen Sie die Genauigkeit Ihrer Aussagen.

   Es sei A die Designmatrix eines linearen Regressionsmodells (Begriff siehe [4]). Berechnen Sie für jedes Ihrer linearen Regressionsmodelle die *Kondition* der Matrix $A^T A$: $\text{cond}(A^T A) = \frac{\max_k(\lambda_k)}{\min_k(\lambda_k)}$. Dabei ist λ_k der k-te Eigenvektor von $A^T A$. Idealerweise ist die Kondition nur wenig größer als 1. Bei „zu großen“ Werten spricht man von einem schlecht konditionierten Problem. Das ist folgendermaßen zu verstehen:

   Der Parametervekor $\beta = (\beta_0, \beta_1)$ bzw. $\beta = (\beta_0, \beta_1, \beta_2)$ wird über Lösung des linearen Gleichungssystems´

   $$A^T A \beta = A^T z \text{ mit dem Messvektor } z = (z_1, ..., z_n)$$

   bestimmt. Das System wird in Mathematica – und jedem anderen Tool – in der Regel numerisch gelöst, wobei Rechenfehler entstehen. Es bezeichne β die tatsächliche Lösung und $\tilde{\beta}$ die numerische bestimmte Lösung. Weiter sei ϵ die relative Genauigkeit der Messung. Für den relativen Fehler bei Berechnung von β gilt:

   $$\frac{\left\|\tilde{\beta} - \beta\right\|}{\|\beta\|} \lessapprox \text{cond}(A^T A)\epsilon$$

Als Faustregel gilt, dass log(cond(A) Dezimalstellen durch Rundungsfehler verloren gehen. Auf wie viele Stellen genau habe Sie die Parameter β_k betimmt? Argumentieren Sie mit der Kondition der Matrix. Ist β für die gestellte Aufgabe genau genug bestimmt, wenn Mathematica (bzw. das Moduel aus der Python-Bibliothek) naiv das Gauß-Verfahren verwendet?

In Wirklichkeit nutzt Mathematica (bzw. das Modul aus der Python-Bibliothek) die QR-Zerlegung. Warum ist das Problem nun viel besser konditioniert?

2.3 Vergleich mit einem Experiment von 2012

2012 fand dieses Experiment schon bereits einmal statt. Die Messergebnisse von 2012 werden Ihnen zur Verfügung gestellt. Vergleichen Sie die Ergebnisse. Legen Sie dabei besonderen Fokus auf die Abhängigkeit von der Wunschgeschwindigkeit.

1. Vergleichen Sie die Regressionen von 2012 und 2017.
2. Vergleichen Sie die Teststatistiken.
3. Fassen Sie die Daten zusammen und erstellen Sie eine Gesamtstatistik.

2.4 Interpretation der Ergebnisse

Nehmen Sie jeweils zu Ihren Ergebnissen Stellung. Welche Probleme, welche Einschränkungen haben Sie bei Ihren Untersuchungen entdeckt? Welches Modell würden Sie empfehlen?

3 Teameinteilung für die Bearbeitung der Aufgaben

Sie arbeiten in Dreier- bzw. Viererteams.

Literatur

[1] M. Davidich and G. Köster. Towards automatic and robust adjustment of human behavioral parameters in a pedestrian stream model to measured data. *Safety Science*, 50:1253–1260, 2012.

[2] L. Fahrmeir, T. Kneib, and S. Lang. *Regression – Modelle, Methoden und Anwendungen*. Springer, 2009.

[3] W. A. Stahel. *Statistische Datenanalyse – Eine Einführung für Naturwissenschaftler*. Vieweg, 2002.

[4] G. Strang. *Linear Algebra*.

[5] U. Weidmann. Transporttechnik für Fussgänger. *Schriftenreihe des IVT*, 90, 1992.