R/exams: A One-for-All Exams Generator
Written Exams, Online Tests, and Live Quizzes with R

Achim Zeileis

http://www.R-exams.org/
R/exams: A One-for-All Exams Generator
Written Exams, Online Tests, and Live Quizzes with R

Achim Zeileis

http://www.R-exams.org/
R/exams: A One-for-All Exams Generator
Written Exams, Online Tests, and Live Quizzes with R

Achim Zeileis

http://www.R-exams.org/
R/exams: A One-for-All Exams Generator
Written Exams, Online Tests, and Live Quizzes with R

Achim Zeileis

http://www.R-exams.org/
Motivation and challenges

Motivation:

• Many of us teach large lecture courses, also as support for other fields.
• For example, statistics, probability, or mathematics in curricula such as business and economics, social sciences, psychology, etc.
• At WU Wien and Universität Innsbruck: Some courses are attended by more than 1,000 students per semester.
• Several lecturers teach lectures and tutorials in parallel.

Strategy:

• Individualized organization of learning, feedback, and assessment.
• The same pool of exercises at the core of all parts of the course.
<table>
<thead>
<tr>
<th></th>
<th>Learning</th>
<th>Feedback</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Synchronous</strong></td>
<td>Lecture</td>
<td>Live quiz</td>
<td>Written exam</td>
</tr>
<tr>
<td></td>
<td>Live stream</td>
<td>(+ Tutorial)</td>
<td></td>
</tr>
<tr>
<td><strong>Asynchronous</strong></td>
<td>Textbook</td>
<td>Self test</td>
<td>Online test</td>
</tr>
<tr>
<td></td>
<td>Screencast</td>
<td>(+ Forum)</td>
<td></td>
</tr>
</tbody>
</table>
## Motivation and challenges

<table>
<thead>
<tr>
<th>Learning</th>
<th>Feedback</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Synchronous</strong></td>
<td>Lecture</td>
<td>Live quiz</td>
</tr>
<tr>
<td></td>
<td>Live stream</td>
<td>(+ Tutorial)</td>
</tr>
<tr>
<td><strong>Asynchronous</strong></td>
<td>Textbook</td>
<td>Self test</td>
</tr>
<tr>
<td></td>
<td>Screencast</td>
<td>(+ Forum)</td>
</tr>
</tbody>
</table>

**Learning:**

- *Standard:* Textbook along with presentation slides.
- *Streaming:* Videos streamed simultaneously or (pre-)recorded.
## Motivation and challenges

<table>
<thead>
<tr>
<th></th>
<th>Learning</th>
<th>Feedback</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Synchronous</strong></td>
<td>Lecture</td>
<td>Live quiz</td>
<td>Written exam</td>
</tr>
<tr>
<td></td>
<td>Live stream</td>
<td>(+ Tutorial)</td>
<td></td>
</tr>
<tr>
<td><strong>Asynchronous</strong></td>
<td>Textbook</td>
<td>Self test</td>
<td>Online test</td>
</tr>
<tr>
<td></td>
<td>Screencast</td>
<td>(+ Forum)</td>
<td></td>
</tr>
</tbody>
</table>

**Feedback & assessment:**

- *Scalability*: Randomized dynamic exercises required.
- *Feedback*: Support for complete correct solutions.
- *Flexibility*: Automatic rendering into different assessment formats.
R package exams

Exercises:
- Each exercise is a single file (either .Rmd or .Rnw).
- Contains question and (optionally) the corresponding solution.
- Dynamic templates if R code is used for randomization.

Answer types:
- Single choice and multiple choice.
- Numeric values.
- Text strings (typically short).
- Combinations of the above (cloze).
R package exams

Output:

- PDF – fully customizable vs. standardized with automatic scanning/evaluation.
- HTML – fully customizable vs. embedded into exchange formats below.
- Moodle XML.
- QTI XML standard (version 1.2 or 2.1), e.g., for OLAT/OpenOLAT.
- ARSnova, TCExam, LOPS, …

Infrastructure: Standing on the shoulders of lots of open-source software…
R package exams

<table>
<thead>
<tr>
<th>Type</th>
<th>Software</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical computing</td>
<td>R</td>
<td>Random data generation, computations</td>
</tr>
<tr>
<td>Writing/reporting</td>
<td>\LaTeX, Markdown</td>
<td>Text formatting, mathematical notation</td>
</tr>
<tr>
<td>Reproducible research</td>
<td>knitr, rmarkdown, Sweave</td>
<td>Dynamically tie everything together</td>
</tr>
<tr>
<td>Document conversion</td>
<td>TtH/TtM, pandoc</td>
<td>Conversion to HTML and beyond</td>
</tr>
<tr>
<td>Image manipulation</td>
<td>ImageMagick, magick, png</td>
<td>Embedding graphics</td>
</tr>
<tr>
<td>Web technologies</td>
<td>base64enc, RCurl,</td>
<td>Embedding supplementary files</td>
</tr>
<tr>
<td>Learning management</td>
<td>Moodle, OpenOLAT, ARSnova,</td>
<td>E-learning infrastructure</td>
</tr>
</tbody>
</table>
Dynamic Exercises
Dynamic exercises

Text file:

1. Random data generation (optional).
2. Question.

Examples:

Multiple-choice knowledge quiz with shuffled answer alternatives.
Which of these are open-source learning management systems?

Dynamic numeric arithmetic exercise.
What is the derivative of $f(x) = x^a e^{b \cdot x}$, evaluated at $x = c$?
Dynamic exercises: .Rmd

Example: Which of these are open-source learning management systems?
Example: Which of these are open-source learning management systems?

Question
=========
Which of these are open-source learning management systems?

Answerlist
----------
* Canvas
* Ilias
* Moodle
* OLAT
* Blackboard
* Desire2Learn
Dynamic exercises: .Rmd

**Example:** Which of these are open-source learning management systems?

Solution
---------
An overview of learning management systems can be found in Wikipedia at <https://en.wikipedia.org/wiki/List_of_learning_management_systems>.

**Answerlist**
---------
* True. Canvas is developed by Instructure Inc. under the Affero General Public License Version 3.
* True. Ilias is an open-source system released under the General Public License Version 3.
* True. Moodle is an open-source system released under the General Public License Version 3.
* True. OLAT is developed by the University of Zurich under the Apache License 2.0 License.
* False. Blackboard Learn is a commercial system developed by Blackboard Inc.
* False. Brightspace is a commercial system developed by Desire2Learn.
Dynamic exercises: .Rmd

Example: Which of these are open-source learning management systems?

Solution
========
An overview of learning management systems can be found in Wikipedia at <https://en.wikipedia.org/wiki/List_of_learning_management_systems>.

Answerlist
---------
* True. Canvas is developed by Instructure Inc. under the Affero General Public License Version 3.
* True. Ilias is an open-source system released under the General Public License.Version 3.
* True. Moodle is an open-source system released under the General Public License Version 3.
* True. OLAT is developed by the University of Zurich under the Apache License 2.0 License.
* False. Blackboard Learn is a commercial system developed by Blackboard Inc.
* False. Brightspace is a commercial system developed by Desire2Learn.

Meta-information
=================
exname: Learning management systems
extype: mchoice
exsolution: 111100
exshuffle: 5
Example: What is the derivative of $f(x) = x^a e^{b \cdot x}$, evaluated at $x = c$?
Example: What is the derivative of $f(x) = x^a e^{b \cdot x}$, evaluated at $x = c$?

```r
<<echo=FALSE, results=hide>>=
## parameters
a <- sample(2:9, 1)
b <- sample(seq(2, 4, 0.1), 1)
c <- sample(seq(0.5, 0.8, 0.01), 1)
## solution
res <- exp(b * c) * (a * c^(a-1) + b * c^a)
@
Example: What is the derivative of $f(x) = x^a e^{b \cdot x}$, evaluated at $x = c$?

```
<<echo=FALSE, results=hide>>=
## parameters
a <- sample(2:9, 1)
b <- sample(seq(2, 4, 0.1), 1)
c <- sample(seq(0.5, 0.8, 0.01), 1)
## solution
res <- exp(b * c) * (a * c^-(a-1) + b * c^a)
@

\begin{question}
What is the derivative of $f(x) = x^{\textcolor{red}{\Sexpr{a}}} e^{\textcolor{red}{\Sexpr{b}}x}$, evaluated at $x = \textcolor{red}{\Sexpr{c}}$?
\end{question}
```
Example: What is the derivative of \( f(x) = x^a e^{b \cdot x} \), evaluated at \( x = c \)?

\begin{solution}
Using the product rule for \( f(x) = g(x) \cdot h(x) \), where
\( g(x) := x^a \) and \( h(x) := e^{b \cdot x} \), we obtain
\begin{eqnarray*}
f'(x) & = & (g(x) \cdot h(x))' = g'(x) \cdot h(x) + g(x) \cdot h'(x) \\
& = & a \cdot x^{a-1} \cdot e^{b \cdot x} + \ldots
\end{eqnarray*}
Evaluated at \( x = c \), the answer is
\[
\left[ e^{b \cdot c} \cdot c^{a-1} \cdot (a + b \cdot c) \right] = \text{fmt}(\text{res}, 6).
\]
Thus, rounded to two digits we have \( f'(c) = \text{fmt}(\text{res}) \).
\end{solution}
Example: What is the derivative of $f(x) = x^a e^{b \cdot x}$, evaluated at $x = c$?

\begin{solution}
Using the product rule for $f(x) = g(x) \cdot h(x)$, where $g(x) := x^{\Sexpr{a}}$ and $h(x) := e^{\Sexpr{b} \cdot x}$, we obtain
\begin{eqnarray*}
f'(x) &=& [g(x) \cdot h(x)]' = g'(x) \cdot h(x) + g(x) \cdot h'(x) \\
& &=& \Sexpr{a} x^{\Sexpr{a} - 1} \cdot e^{\Sexpr{b} \cdot x} + \ldots
\end{eqnarray*}
Evaluated at $x = \Sexpr{c}$, the answer is
\[ e^{\Sexpr{b} \cdot \Sexpr{c}} \cdot \Sexpr{c}^{\Sexpr{a} - 1} \cdot (\Sexpr{a} + \Sexpr{b} \cdot \Sexpr{c}) = \Sexpr{fmt(res, 6)}. \]
Thus, rounded to two digits we have $f'(\Sexpr{c}) = \Sexpr{fmt(res)}$.
\end{solution}

\extype{num}
\exsolution{\Sexpr{fmt(res)}}
\exname{derivative exp}
\extol{0.01}
Dynamic exercises: Single choice

exetype: schoice
exsolution: 010
Dynamic exercises: Single choice

Question
What is the seat of the federal authorities in Switzerland (i.e., the de facto capital)?

(a) Bern
(b) Lausanne
(c) Zurich
(d) St. Gallen
(e) Basel

Knowledge quiz: Shuffled distractors.
Question
What is the derivative of $f(x) = x^3 e^{3.3x}$, evaluated at $x = 0.85$?

(a) 45.97  
(b) 35.82  
(c) 56.45  
(d) 69.32  
(e) 39.31

Numeric exercises: Distractors are random numbers and/or typical arithmetic mistakes.
Dynamic exercises: Multiple choice

```
extype: mchoice
exsolution: 011
```
Question
Which of these are open-source learning management systems?
(a) Ilias
(b) OLAT
(c) Blackboard
(d) Moodle
(e) Canvas

Knowledge quiz: Shuffled true/false statements.
Dynamic exercises: Multiple choice

**Question**
In the following figure the distributions of a variable given by two samples (A and B) are represented by parallel boxplots. Which of the following statements are correct? *(Comment: The statements are either about correct or clearly wrong.)*

- (a) The location of both distributions is about the same.
- (b) Both distributions contain no outliers.
- (c) The spread in sample A is clearly bigger than in B.
- (d) The skewness of both samples is similar.
- (e) Distribution B is about symmetric.

**Interpretations:** Statements that are approximately correct or clearly wrong.
Dynamic exercises: Numeric

extype: num
exsolution: 123.45
Dynamic exercises: Numeric

**Question**
Given the following information:

\[
\begin{align*}
\text{apple} \ + \ \text{pineapple} \ + \ \text{orange} & = 585 \\
\text{banana} \ + \ \text{orange} \ + \ \text{banana} & = 144 \\
\text{apple} \ + \ \text{banana} \ + \ \text{orange} & = 177
\end{align*}
\]

Compute:

\[
\text{banana} \ + \ \text{orange} \ + \ \text{pineapple} = ?
\]

*Numeric exercises: Solving arithmetic problems.*
Dynamic exercises: String

extype: string
exsolution: ANSWER
Dynamic exercises: String

**Question**
What is the name of the R function for Poisson regression?

**Knowledge quiz:** Sample a word/phrase from a given vocabulary or list of question/answer pairs.

exetype: string
exsolution: ANSWER
Dynamic exercises: Cloze

extype: cloze
exclzoetype: mchoice|num
exsolution: 10|123.45
Question
Using the data provided in regression.csv estimate a linear regression of $y$ on $x$ and answer the following questions.

(a) $x$ and $y$ are not significantly correlated / $y$ increases significantly with $x$ / $y$ decreases significantly with $x$

(b) Estimated slope with respect to $x$: 

Exercises with sub-tasks: Several questions based on same problem setting.
One-for-All
One-for-all

- The *same* exercise can be exported into different formats.
- Multiple standalone documents vs. combined exercise pool.
- Multiple-choice and single-choice supported in all output formats.
One-for-All

**Idea:** An exam is simply a list of exercise templates.

```r
R> myexam <- list(
+   "deriv2.Rnw",
+   "fruit2.Rnw",
+   c("ttest.Rnw", "boxplots.Rnw")
+)
```

**Draw random exams:**

- First randomly select one exercise from each list element.
- Generate random numbers/input for each selected exercise.
- Combine all exercises in output file(s) (PDF, HTML, …).
One-for-All

Written exam:

R> exams2nops(myexam, n = 3, dir = odir,
+     language = "nl", institution = "TEA 2018")

Online test:

R> exams2moodle(myexam, n = 10, dir = odir)

Live quiz:

R> exams2arsnova(myexam, n = 1, dir = odir)

Other: exams2pdf(), exams2html(), exams2qti12(), exams2qti21(),...
Written Exams
Written Exams

Flexible: Roll your own.
- Combination with user-specified template in exams2pdf() and exams2pandoc().
- Customizable but typically has to be evaluated “by hand”.

Standardized: “NOPS” format.
- exams2nops() intended for single- and multiple-choice questions.
- Can be scanned and evaluated automatically within R.
- Limited support for open-ended questions that have to be marked by a person.
1. What is the derivative of \( f(x) = x^7e^{3.7x} \), evaluated at \( x = 0.83 \)?

(a) 49.35  
(b) 87.17  
(c) 71.00  
(d) 73.46  
(e) 55.20

2. Given the following information:

\[
\begin{align*}
282 & = a + b + c  \\
137 & = b + c + d  \\
106 & = c + d + e
\end{align*}
\]

Compute:

\[
\begin{align*}
? & = a + b + c + d + e
\end{align*}
\]

(a) 106  
(b) 213  
(c) 161  
(d) 232  
(e) 454

3. The waiting time (in minutes) at the cashier of two supermarket chains with different cashier systems is compared. The following statistical test was performed:

Two Sample t-test

data: Waiting by Supermarket
\( t = -3.3, \) df = 90, \( p \)-value = 1
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
\(-3.227 \text{ Inf}\)
sample estimates:
mean in group Sparag mean in group Consumo
4.045 6.192

Which of the following statements are correct? (Significance level 5%)

(a) The absolute value of the test statistic is larger than 1.96.
(b) A one-sided alternative was tested.
(c) The \( p \)-value is larger than 0.05.
(d) The test shows that the waiting time is longer at Sparag than at Consumo.
(e) The test shows that the waiting time is shorter at Sparag than at Consumo.
Written exams

1. Create
   • As illustrated above.
   • Using `exams2nops()` function, create (individual) PDF files for each examinee.
Written exams

1. Create
   - As illustrated above.
   - Using `exams2nops()`, create (individual) PDF files for each examinee.

2. Print
   - Print the PDF exams, e.g., on a standard printer.
   - …or for large exams at a print shop.
Written exams

3. Exam
- Conduct the exam as usual.
- Collect the completed exams sheets.
Written exams

4. Scan

- Scan exam sheets, e.g., on a photocopier.
- Using `nops_scan()`, process the scanned exam sheets to machine-readable content.
4. Scan

- Scan exam sheets, e.g., on a photocopier.
- Using `nops_scan()`, process the scanned exam sheets to machine-readable content.

5. Evaluate

- Using `nops_eval()`, evaluate the exam to obtain marks, points, etc. and individual HTML reports for each examinee.
- Required files: Correct answers (1.), scans (4.), and a participant list in CSV format.
Written exams

A vizsga eredménye
Név: Jane Doe
Regisztrációs szám: 1501090
Érdemjegy: 5
Pontok: 3.16666666666667

Értékelés

<table>
<thead>
<tr>
<th>Kérdés</th>
<th>Pontok</th>
<th>Adott válasz</th>
<th>Helyes válasz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000000</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>2</td>
<td>0.5000000</td>
<td>abc e</td>
<td>abc__</td>
</tr>
<tr>
<td>3</td>
<td>0.0000000</td>
<td>ab d</td>
<td>ab d</td>
</tr>
<tr>
<td>4</td>
<td>1.0000000</td>
<td>c</td>
<td>bc__</td>
</tr>
<tr>
<td>5</td>
<td>0.6666667</td>
<td>d</td>
<td>ab d</td>
</tr>
<tr>
<td>6</td>
<td>0.0000000</td>
<td>bc e</td>
<td>a c</td>
</tr>
</tbody>
</table>

Vizsgalap

R University
Exam 2015-07-29

A vizsga eredménye
Név: Ambi Dexter
Regisztrációs szám: 9901071
Érdemjegy: 5
Pontok: 1.5

Értékelés

<table>
<thead>
<tr>
<th>Kérdés</th>
<th>Pontok</th>
<th>Adott válasz</th>
<th>Helyes válasz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>a c</td>
<td>d</td>
</tr>
<tr>
<td>2</td>
<td>0.0</td>
<td>a c d e</td>
<td>ab d</td>
</tr>
<tr>
<td>3</td>
<td>0.0</td>
<td>b</td>
<td>e</td>
</tr>
<tr>
<td>4</td>
<td>0.0</td>
<td>c</td>
<td>a c d</td>
</tr>
<tr>
<td>5</td>
<td>0.0</td>
<td>c</td>
<td>b c</td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>abc</td>
<td>a</td>
</tr>
</tbody>
</table>

Vizsgalap

Universität Innsbruck
Klausur 2015-07-29

Personal Data

Family Name: Doe
Given Name: Jane
Signature: [Signature]

Regisztr.

| 1 | 5 | 0 |

Persönliche Daten

Nachname: Dexter
Vorname: Ambi
Unterschrift: [Signature]

Matr.

| 9 | 9 | 1 |
E-Learning

1. Goal

- Online tests with flexible exercise types.
- Possibly: Dynamic supplements and/or complete correct solution.
- Random variations of similar exercises to reduce the risk of cheating.
- Use university’s learning management system, e.g., Moodle, …
2. Create

- Draw random replications from exercise templates, e.g., via `exams2moodle()`, …
- Automatically embed these into exchange file format (typically via HTML/XML).
E-Learning

2. Create

- Draw random replications from exercise templates, e.g., via `exams2moodle()`, ...
- Automatically embed these into exchange file format (typically via HTML/XML).

3. Import

- Import in learning management system.
- From there handling “as usual” in the system.
E-Learning: Online test

Preview question: R01 Q1 : deriv

What is the derivative of $f(x) = x^3 \cdot 3.3^x$, evaluated at $x = 0.75$?

Answer: 51.83594

Check

Using the product rule for $f(x) = g(x) \cdot h(x)$, where $g(x) = x^3$ and $h(x) = 3.3^x$, we obtain

$$f'(x) = (g'(x) \cdot h(x)) + (g(x) \cdot h'(x))$$

$$= 3x^2 \cdot 3.3^x + x^3 \cdot 3.3 \cdot (3x^2)$$

$$= e^{3.3x} - 3x^2 \cdot (3 + 3.3x).$$

Evaluated at $x = 0.75$, the answer is

$$e^{3.3 \cdot 0.75} \cdot 0.75^2 \cdot (3 + 3.3 \cdot 0.75) = 36.501945.$$ 

Thus, rounded to two digits we have $f'(0.75) = 36.59$.

The correct answer is: 36.59

Start again  Save  Fill in correct responses  Submit and finish  Close preview

Preview question: R01 Q6 : Im

Using the data provided in regression.csv estimate a linear regression of y on x and answer the following questions.

- a. x and y are not significantly correlated
- b. Estimated slope with respect to x: -0.08

Check

b = -0.087
p = 0.051
E-Learning: Online test

Two Sample t-test

data:  Waiting by Supermarket

k = -0.09169, df = 53, p-value = 0.3084
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:

Inf -0.0450572

sample estimates:
mean in group Sparag mean in group Consumo
7.68248   7.662992

Which of the following statements are correct? (Significance level 5%)

- a. The absolute value of the test statistic is larger than 1.96.
- b. A one-sided alternative was tested.
- c. The p-value is larger than 0.05.
- d. The test shows that the waiting time is longer at Sparag than at Consumo.

Using the data provided in regression.csv estimate a linear regression of y on x and answer the following questions.

- a.1. x and y are not significantly correlated
- a.2. y increases significantly with x
- a.3. y decreases significantly with x

b. Estimated slope with respect to x: 0.08
E-Learning: Live quiz

Which of these institutions already hosted a useR! or eRum conference?

- Universität Wien
- ETH Zürich
- Københavns Universitet

What is the derivative of \( f(x) = x^9 e^{2x} \), evaluated at \( x = 0.7 \)?

- 2.43
- 3.70
- 2.10

Given the following information:

\[
\begin{align*}
\text{470} & = + + \\
\text{502} & = + + \\
\text{166} & = + + \\
\end{align*}
\]

Compute:

\[
\begin{align*}
\text{?} & = + + \\
\end{align*}
\]
What Else?
What else?

Under development:

- Many volunteers: Internationalization for “NOPS” exams.
- Nikolaus Umlauf: Exercise “stress tester”.
- Nikolaus Umlauf: Graphical exams manager based on shiny that can be used on a local machine or on a server.
- Achim Zeileis: Reports for lecturers based on IRT models.
- Niels Smits: Better management of exercise categories.
- Niels Smits, Claus Ekstrøm, Nikolaus Umlauf: Canvas interface based on QTI 1.2.
- Mirko Birbaumer, Andreas Melillo, Achim Zeileis: Ilias interface based on QTI 1.2.
Please mark the boxes carefully: ❌ Not marked: □ or ■
This document is scanned automatically. Please keep clean and do not bend or fold.
Only clearly marked and positionally accurate crosses will be processed!

Answers 1 - 15
1 □ □ □ □ □
2 □ □ □ □ □

Answers 16 - 21
16 □ □ □ □ □
17 □ □ □ □ □

Merci de cocher soigneusement: ❌ Non coché: □ ou ■
Cet examen sera corrigé par un système automatisé. Ne pas plier, corner ni tacher. Merci d’utiliser un stylo à bille bleu ou noir.
Seul les marques lisibles et bien positionnées seront évaluées.

Réponses 1 - 15
1 □ □ □ □ □
2 □ □ □ □ □

Réponses 16 - 21
16 □ □ □ □ □
17 □ □ □ □ □

A válaszát jelölje egyértelmű x-é! ❌ Jelöllett cella: □ vagy □
A vizsgalap szkennelése automatikusan történik, ezért kérjük, hogy kék vagy fekete tollat.
Kizárólag az egyértelmű és pontosan megjelölt válaszok képesek reagálni.

Válaszok 1 - 15
1 □ □ □ □ □
2 □ □ □ □ □

Válaszok 16 - 21
16 □ □ □ □ □
17 □ □ □ □ □

Jensen, Messner
Zeileis
Kogelnik
Nordhausen
Allignol
Stauffer
Juraić, Kecojevic
Daróczy, Tóth
Zambella
Smits
Calvão, Dellinger,
Petutschnig (pt-PT/pt-BR)
Gatu
Demeshev
Fabsic
Kecojevic
Er

More contributions welcome …
R> s <- stresstest_exercise("deriv2.Rnw")
R> plot(s)
Stress tester

R> s <- stresstest_exercise("deriv2.Rnw")
R> plot(s)
Graphical exams manager

Question

In the following figure the distributions of a variable given by two samples (A and B) are represented by parallel boxplots. Which of the following statements are correct? (Comment: The statements are either all correct or clearly wrong.)

Solution

a. The location of both distributions is about the same.
b. Both distributions contain no outliers.
c. The spread in sample A is clearly bigger than in B.
d. The skewness of both samples is similar.
e. Distribution B is left-skewed.

a. True. Both distributions have a similar location.
b. True. Both distributions have no observations which deviate more than 1.5 times the interquartile range from the box.
Examining exams

**Report:** Exercise difficulty, student performance, unidimensionality, fairness.

**Methods:** Psychometrics, especially item response theory.

**Example:** End-term exam from first-year mathematics course for business and economics students at Universität Innsbruck.

- 729 students (out of 941 registered).
- 13 single-choice exercises on the basics of analysis, linear algebra, financial mathematics.
- Two groups with partially different pools of exercise templates.

```r
R> library("psychotools")
R> data("MathExam14W", package = "psychotools")
R> mex <- subset(MathExam14W, nsolved > 0 & nsolved < 13)
```
Examining exams

**Item difficulty**: Raw proportions vs. Rasch model.

R> plot(mex$solved, ...)
R> mr <- raschmodel(mex$solved)
R> plot(mr, ...)

[Graph showing item difficulty parameters with bars and a line chart.]

---

42
Examining exams

**Student performance:** Points and person-item map.

```r
R> hist(MathExam14W$points, ...)
R> piplot(mr)
```

![Histogram and person-item map](image)
Examining exams

**Unidimensionality:** Principal component analysis.

R> pr <- prcomp(mex$solved, scale = TRUE)
R> plot(pr, ...)
R> biplot(pr, ...)
Examining exams

**Fairness:** Differential item functioning.

R> ma <- anchortest(solved ~ group, data = mex, adjust = "single-step")
R> plot(ma$final_tests, ...)

---

**Item difficulty parameters**

-2 -1 0 1 2 3

-2 -1 0 1 2 3

Group 1  Group 2

---

**Family-wise 95% confidence intervals**

-2 -1 0 1

---

**Legend:**
- quad
- deriv
- elasticity
- integral
- interest
- annuity
- payflow
- matrix
- planning
- hesse
- implicit
- lagrange

---

**Graph:**

- Item difficulty parameters
- Family-wise 95% confidence intervals
Recommendations

If you want to try 📚 R/exams:

• Start with simple exercises before moving to more complex tasks.
• Focus on content of exercises.
• Don’t worry about layout/formatting too much.
• Try to build a team (with lecturers, assistants, etc.).
• Use exercise types creatively.
• Don’t be afraid to try stuff, especially in formative assessments.
• Thorough quality control for dynamic exercises before summative assessments.
Resources

Contributors: Zeileis, Grün, Leisch, Umlauf, Smits, Birbaumer, Ernst, Keller, Krimm, Stauffer.

Links:
- CRAN  https://CRAN.R-project.org/package=exams
- Forum  http://R-Forge.R-project.org/forum/?group_id=1337
- StackOverflow  https://stackoverflow.com/questions/tagged/exams
- Twitter  @AchimZeileis

References: