



# R/exams: A One-for-All Exams Generator

Achim Zeileis

https://www.R-exams.org/

### Goals:

- Enable individualized organization of learning, feedback, and assessment.
- Foster continuing active participation.
- Encourage learning from peers and supporting each other.

### Goals:

- Enable individualized organization of learning, feedback, and assessment.
- Foster continuing active participation.
- Encourage learning from peers and supporting each other.

### **Challenges:**

- Deal with large(-ish) student groups.
- Discourage cheating without learning.

### Strategy:

- High-frequency low-stakes formative assessment in supportive environment.
- Strict summative assessment in controlled environment.
- One pool of exercises as the basis for all assessments.

### Strategy:

- High-frequency low-stakes formative assessment in supportive environment.
- Strict summative assessment in controlled environment.
- One pool of exercises as the basis for all assessments.

Technology: R/exams software & learning management systems.

- Individualization: Randomized dynamic exercise pools.
- Feedback: Support for complete correct solutions.
- Automatic evaluation: Rendering into different (closed) assessment formats.

### Example: Formative assessments

### Some possibilities:

- Quick quiz: Recap of knowledge from previous session or preparation.
- Asynchronous test: Several days, encourage group work, support in forum.
- *Synchronous test*: In learning groups in live sessions (possibly completion of individual tests afterwards).
- Open-ended tasks: Upload of solutions online, discussion in live sessions.

# Example: First-year mathematics

Structure: Mandatory.

- 2-hour lecture (VO), 500+ participants.
- 2-hour tutorial (PS), up to 40 participants per tutorial.

# Example: First-year mathematics

### Structure: Mandatory.

- 2-hour lecture (VO), 500+ participants.
- 2-hour tutorial (PS), up to 40 participants per tutorial.

### Weekly schedule:

- Learning: Textbook (soon online), screencasts, slides.
- Flipped classroom: Overview, questions & answers in lecture session.
- Formative assessment: Numeric online test, several days, support in forum.
- Feedback: Discussion of more complex exercises in tutorial session.
- Formative assessment: Another online test in tutorial learning groups.

# Example: First-year mathematics

### Structure: Mandatory.

- 2-hour lecture (VO), 500+ participants.
- 2-hour tutorial (PS), up to 40 participants per tutorial.

### Weekly schedule:

- Learning: Textbook (soon online), screencasts, slides.
- Flipped classroom: Overview, questions & answers in lecture session.
- Formative assessment: Numeric online test, several days, support in forum.
- Feedback: Discussion of more complex exercises in tutorial session.
- Formative assessment: Another online test in tutorial learning groups.

Summative assessment: Written single-choice exams (mid-term & end-term).

# Example: Data analytics

Structure: Elective.

- 2-hour lecture (VU), 40–60 participants.
- 1-hour tutorial (VU), 15–40 participants per tutorial.

# Example: Data analytics

Structure: Elective.

- 2-hour lecture (VU), 40–60 participants.
- 1-hour tutorial (VU), 15–40 participants per tutorial.

Weekly schedule:

- *Learning:* Lecture session, online textbook/slides.
- Formative assessment: Quiz & numeric test, several days, support in forum.
- Formative assessment: Open-ended practical task, several days, file upload.
- Feedback: Discussion of practical tasks in tutorial session.

# Example: Data analytics

Structure: Elective.

- 2-hour lecture (VU), 40–60 participants.
- 1-hour tutorial (VU), 15–40 participants per tutorial.

Weekly schedule:

- Learning: Lecture session, online textbook/slides.
- Formative assessment: Quiz & numeric test, several days, support in forum.
- Formative assessment: Open-ended practical task, several days, file upload.
- Feedback: Discussion of practical tasks in tutorial session.

**Summative assessment:** Online exams with open-ended theory questions and individualized practical tasks (mid-term & end-term).

# 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 (either closed/short or open-ended).
- 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 Canvas, OpenOlat, ILIAS.
- Blackboard (partially based on QTI 1.2)
- learnr, Particify, LOPS, ...

Infrastructure: Standing on the shoulders of lots of open-source software...

# R package exams

Туре	Software	Purpose
Statistical computing	R	Random data generation, computations
Writing/reporting	Markdown, ᡌ᠋TEX	Text formatting, mathematical notation
Reproducible research	knitr, rmarkdown, Sweave	Dynamically tie everything together
Document conversion	pandoc, TtH/TtM	Conversion to HTML and beyond
Image manipulation	lmageMagick, magick, png	Embedding graphics
Web technologies	base64enc, RCurl,	Embedding supplementary files
Learning management	Moodle, OpenOlat, Canvas, Particify,	E-learning infrastructure

# **Dynamic Exercises**



# Dynamic exercises

### Text file:

- 1 Random data generation (optional).
- Question.
- **3** Solution (optional).
- 4 Metainformation.

### **Examples:**



Multiple-choice knowledge quiz with shuffled answer alternatives. Which of the following cities are the capital of the corresponding country?



Dynamic numeric arithmetic exercise.

### Dynamic exercises: .Rmd

**Example:** Which of the following cities are the capital of the corresponding country?

### Dynamic exercises: . Rmd

# **Example:** Which of the following cities are the capital of the corresponding country?

Question \_\_\_\_\_ Which of the following cities are the capital of the corresponding country? Answerlist \* Lagos (Nigeria) \* São Paulo (Brazil) \* Toronto (Canada) \* Auckland (New Zealand) \* Istanbul (Turkey) \* Zürich (Switzerland) \* Tokyo (Japan) \* New Delhi (India) \* Astana (Kazakhstan) \* Warsaw (Poland) \* Rivadh (Saudi Arabia)

### Dynamic exercises: . Rmd

**Example:** Which of the following cities are the capital of the corresponding country?

So ==	lution
An	swerlist
- * * * * * * * * * *	False. The capital of Nigeria is Abuja. False. The capital of Brazil is Brasilia. False. The capital of Canada is Ottawa. False. The capital of New Zealand is Wellington. False. The capital of Turkey is Ankara. False. The de facto capital of Switzerland is Bern. True. Tokyo is the capital of Japan. True. New Delhi is the capital of India. True. Astana is the capital of Kazakhstan. True. Warsaw is the capital of Poland.
*	True. Riyadh is the capital of Saudi Arabia.

### Dynamic exercises: . Rmd

**Example:** Which of the following cities are the capital of the corresponding country?

```
'```{r data generation, 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)
'```</pre>
```

**Example:** What is the derivative of  $f(x) = x^a e^{b \cdot x}$ , evaluated at x = c?

```
'```{r data generation, 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)
'```</pre>
```

Question

\_\_\_\_\_

What is the derivative of  $f(x) = x^{r} a^{ t} a^{ t} a^{ t}$ , evaluated at  $x = r c^{ t}$ ?

```
<<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 $f(x) = x^{\Sexpr{a}} e^{\Sexpr{b}x}$,
evaluated at $x = \Sexpr{c}?
\end{question}
```

```
Solution
_____
Using the product rule for f(x) = g(x) \setminus dt h(x), where
g(x) := x^{r} a^{s} and h(x) := e^{r} b^{s} x^{s}, we obtain
$$
\begin{aligned}
f'(x) &= [g(x) \setminus dot h(x)]' = g'(x) \setminus dot h(x) + g(x) \setminus dot h'(x) \setminus
\end{aligned}
$$
Meta-information
_____
extype: num
exsolution: r fmt(res)
exname: derivative exp
extol: 0.01
```

```
\begin{solution}
 Using the product rule for f(x) = g(x) \cdot dot h(x), where
 g(x) := x^{\sqrt{b}x}, we obtain
 \begin{eqnarray*}
f'(x) \& = \& [g(x) \setminus cdot h(x)]' = g'(x) \setminus cdot h(x) + g(x) \setminus cdot h'(x) \setminus b(x) = \& [g(x) \setminus cdot h(x)]' = g'(x) \setminus cdot h(x) + g(x) \setminus cdot h(x) = g'(x) \cap cdot h(x) = 
   \end{eqnarray*}
   \end{solution}
   \extvpe{num}
   \exsolution{\Sexpr{fmt(res)}}
   \exname{derivative exp}
   \left( 0, 01 \right)
```

### Dynamic exercises: Single choice



extype: schoice
exsolution: 010

# Dynamic exercises: Single choice



extype: schoice exsolution: 010

<ul> <li>(a) Basel</li> <li>(b) Bern</li> <li>(c) Zurich</li> <li>(d) Geneva</li> </ul>	
<ul><li>(b) Bern</li><li>(c) Zurich</li><li>(d) Geneva</li></ul>	
(c) Zurich (d) Geneva	
(d) Geneva	
(d) donota	
(e) Lausanne	

Knowledge quiz: Shuffled distractors.

# Dynamic exercises: Single choice



extype: schoice exsolution: 010

<b>Question</b> What is the derivative of $f(x) = x^9 e^{3.9x}$ , evaluated at $x = 0.82$ ?
(a) 61.05
(b) 49.15
(c) 72.53
(d) 45.04
(e) 61.47

*Numeric exercises:* Distractors are random numbers and/or typical arithmetic mistakes.

# Dynamic exercises: Multiple choice

×	==

extype: mchoice
exsolution: 011

# Dynamic exercises: Multiple choice



extype: mchoice exsolution: 011

# Question Which of the following cities are the capital of the corresponding country? (a) New Delhi (India) (b) Tokyo (Japan) (c) Lagos (Nigeria) (d) Auckland (New Zealand) (e) Astana (Kazakhstan)

Knowledge quiz: Shuffled true/false statements.

# Dynamic exercises: Multiple choice



extype: mchoice exsolution: 011

### 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.)



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

### Dynamic exercises: Numeric



extype: num exsolution: 123.45

# Dynamic exercises: Numeric



extype: num exsolution: 123.45



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 extracting the fitted loglikelihood from a fitted (generalized) linear model object?

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

extype: string exsolution: ANSWER

# Dynamic exercises: String



extype: string exsolution: ANSWER

Question Consider the following regression results:
Call:
lm(formula = log(y) - log(x), data = d)
Residuals:
Min 1Q Median 3Q Max
-5.490 -1.056 0.102 1.593 5.187
Coefficients:
Estimate Std. Error t value Pr(> t )
(Intercept) -0.039 0.304 -0.13 0.90
log(x) 0.217 0.308 0.70 0.48
Residual standard error: 2.24 on 53 degrees of freedon Multiple R-squared: 0.00927, Adjusted R-square F-statistic: 0.496 on 1 and 53 DF, p-value: 0.484
Describe how the response y depends on the regressor x.

Open-ended question: Answer in an essay editor and/or by file upload (via exstringtype).

### Dynamic exercises: Cloze



extype: cloze
exclozetype: mchoice|num
exsolution: 10|123.45

# Dynamic exercises: Cloze



### 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.

extype: cloze
exclozetype: mchoice|num
exsolution: 10|123.45

# **One-for-All**



All the stand the stand stand

Million management

Alasi, Tana and Angel and

Teo Sample t-test

 The existing time (in minutes) at the cashiar of two supermarket choice with dele systems is compared. The following statistical test was performed.

Family 17091900001

# 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> myexam <- list(
+ "capitals.Rmd",
+ "deriv2.Rmd",
+ 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

### **Online test:**

```
R> exams2openolat(myexam, n = 10, dir = odir)
```

Live quiz:

```
R> exams2particify(myexam, n = 1, dir = odir)
```

### Written exam:

```
R> exams2nops(myexam, n = 3, dir = odir,
+ language = "de", date = "2023-03-29",
+ institution = "Universität Innsbruck", logo = "uibk-logo-bw.png")
```

**Other:** exams2pdf(), exams2html(), exams2moodle(), exams2canvas(), ...





### 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., OpenOlat, Moodle, ...

### Scenarios:

- Short quizzes conducted in-class.
- Online tests conducted over several days.
- E-exams conducted in-class or remotely.



### 2. Create

- Draw random replications from exercise templates, e.g., via exams2openolat(), ...
- Automatically embed these into exchange file format (typically via HTML/XML).



# TARSnova The second

### 2. Create

- Draw random replications from exercise templates, e.g., via exams2openolat(), ...
- 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



# E-Learning: Online test

OpenOLAT - infinite learning - Mozilla Firefox	×
ile <u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp	
• OpenOLAT - infinite learn × +	
$\rightarrow$ $\mathcal{C}$ $\widehat{\mathbf{a}}$ $\bigcirc$ $\widehat{\mathbf{b}}$ https://lms-t.uibk.ac.at/auth/Repo $\square$ $\square$ $\square$ $\bigcirc$ $\widehat{\mathbf{c}}$ $\bigcirc$ $\checkmark$ $\checkmark$ $\checkmark$	≡
🕽 DuckDuckGo G Search G Images 🛞 OpenStreetMap 🎇 Maps   🎇 EO w ikipedia 🔅	»
eRum-2018  Show description Show description	
Question 1 point • Not answered	
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	
<pre>data: Wailing by Supermarkst t = 0.50510, of = 135, p-value = 0.3084 alternative hypothesis: rue difference in means is less than 0 95 percent confidence interval:</pre>	
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.	ĺ
$\ensuremath{\boxtimes}$ 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-Learning: Live quiz

		8	* 🗑 🗟 .	1 69%	02:06
☆ â		ars.uibk.ac.a	t	$( \!$	Q
	1	2	3		
Which tal of	n of the f the corr	following citi esponding c	es are th ountry?	ne cap	)i-
	Lagos	(Nigeria)			
	Istanb	ul (Turkey)			
	Zürich	(Switzerland	I)		
	Warsa	w (Poland)			
	New D	elhi (India)			
<b>A</b> Overvi	ew	A&Q		R/exa	ns
<	>	<u>م</u>	2	ה	≡

	5	1 <b>* 10</b> 🖘 (	02:06 🛢 %9%
☆ 🕯	ars.uibk.ac.	at (	J (J
1	2	3	
What is the evaluated at	derivative of $x=0.7?$	$f(x) = x^8$	$e^{3.1x}$ ,
6.28			
0 7.34			
5.77			
0 4.34			
_			
ft Overview	<b>P</b> Q&A	R	/exams
$\langle \rangle$	<u>۵</u> ۲	<u>}</u> ∂	Ξ





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.



### Exam: 23032900001 1. Which of the following cities are the capital of the corresponding country? (a) New Dalhi (India) (b) Tokyo (Japan) (c) Lagos (Nigeria) (d) Auckland (New Zealand) (a) Astena (Kazakhatan) 2. What is the derivative of $f(x) = x^{q} e^{2.3x}$ , evaluated at x = 0.897(a) 19.60 (b) 27.40 (c) 30.45 (d) 31.09 (a) 34.42 3. In the following fours the distributions of a variable given by two samples (A and B) are In the introduct of a variable provided the interview of a variable given by two samples (e and a) are represented by parallel boxplots. Which of the following statements are correct? (Comment: The statements are either above correct or clearly around ). 25 .... (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.



### 1. Create

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



### 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.



### 3. Exam

- Conduct the exam as usual.
- Collect the completed exams sheets.



### 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.

### A vizsga eredménye

Név: Jane Doe Regisztrációs szám: 1501090 Érdemiegy: 5 Pontok: 3.16666666666667

### Értékelés

Kérdés	Pontok	Adott válasz	Helyes válasz
1	1.0000000	c	c
2	0.5000000	abc_e	abc
3	0.0000000		ab_d_
4	1.0000000	C	_bc
5	0.6666667	d_	ab_d_
6	0.0000000	_bc_e	a_c

Vizsgalap

### **R** University +Exam 2015-07-29 Personal Data Registra Family Name: DOE 1.5.0 0 Given Name: JANE 1 🛛 🗆 Signature: 2 🗆

### A vizsga eredménye

Név:	Ambi Dexter
Regisztrációs szám:	9901071
Érdemjegy:	5
Pontok:	1.5

### Értékelés

Kérdés	Pontok	Adott válasz	Helyes válasz
1	0.0	a_c	d_
2	0.0	a_cde	ab_d_
3	0.0	_b	e
4	0.0		a_cd_
5	0.0		_bc
6	1.5	abc	a

### Vizsgalap

### Universität Innsbruck +

Klausur 2015-07-29





### 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.).
- Connect to experienced users.
- Use exercise types creatively.
- Don't be afraid to try stuff, especially in formative assessments.
- Thorough quality control for dynamic exercises before summative assessments.

### Installation:

- 1 R, Rtools on Windows, RStudio recommended for beginners.
- **2** R package *exams* (including dependencies).
- **③** LATEX for producing PDF output.
- Pandoc (e.g., provided along with RStudio).
- **6** Possibly further tools needed for scanning NOPS exams.

More details: https://www.R-exams.org/tutorials/installation/

First steps: Create exams skeleton.

R> exams\_skeleton()

**Output:** 

- demo-\*.R scripts.
- exercises/ folder with all .Rmd/.Rnw exercises.
- templates/ folder with various customizable templates.
- nops/ folder (empty) for exams2nops() output.

More details: https://www.R-exams.org/tutorials/first\_steps/

Quality control: Stress testing.

- Generate a large number of random versions of an exercise.
- Check for errors, warnings, long computation times, ...
- Especially for numeric exercises: Check solution distribution, outliers, dependency on randomized parameters.
- Especially for multiple-choice exercises: Check shuffling of correct answers.

More details: https://www.R-exams.org/tutorials/stresstest/

### Resources

### **Contributors:**

Zeileis, Grün, Leisch, Umlauf, Smits, Birbaumer, Ernst, Keller, Krimm, Sato, Stauffer, Wickelmaier.

### Links:

Web	https://www.R-exams.org/
CRAN	https://CRAN.R-project.org/package=exams
StackOverflow	https://stackoverflow.com/questions/tagged/r-exams
Forum	https://R-Forge.R-project.org/forum/?group_id=1337
Mastodon	@zeileis@fosstodon.org
Twitter	@AchimZeileis