

R/EXAMS JAKO WSPARCIE W WERYFIKACJI WIEDZY STUDENTÓW

Agnieszka WAŁACHOWSKA

Politechnika Gdańska, Wydział Zarządzania i Ekonomii
tel.: 58 347 13 21 e-mail: agnwalac@pg.edu.pl

Streszczenie: Przydatność zasobów e-learningowych, takich jak testy i quizy online, można dostrzec w wielu sytuacjach: duża liczba studentów uczęszczających na dane zajęcia, utrwalanie, podsumowanie i weryfikacja wiedzy czy nauczanie na odległość. Celem artykułu jest zaprezentowanie pakietu R/exams dostępnego w programie statystycznym R jako wsparcia do tworzenia losowych i dynamicznych pytań oraz ich wykorzystania na platformie Moodle. Omówione zostaną różne typy pytań na przykładzie zadań ze statystyki.

Słowa kluczowe: egzamin, test, pakiet „exams”, platforma Moodle.

1. WPROWADZENIE

1.1. Motywacja

W związku z pandemią COVID-19 uczelnie zostały zmuszone do nagłego przejścia z nauczania tradycyjnego na formę zdalną. Konieczne było opracowanie nowych metod przekazywania studentom wiedzy na odległość. Nastąpił wzrost zainteresowania oprogramowaniem do wideokonferencji, a gdy okazało się, że w formie zdalnej mają zostać przeprowadzone również egzaminy, dodatkowo wzrosło zainteresowanie narzędziami, które pozwolą na ograniczenie możliwości ściągania i sposobu na przeprowadzenie uczciwego zaliczenia przedmiotów. Jedną z takich możliwości daje platforma e-learningowa Moodle poprzez przeprowadzenie egzaminów w formie quizów. Nadal jednak pozostawał problem w stworzeniu odpowiedniej ilości pytań na podobnym poziomie trudności, ale zmodyfikowanych na tyle, aby uniemożliwić studentom wymianę swoich rozwiązań. W przypadku prowadzonych przez autorkę przedmiotów związanych z rachunkiem prawdopodobieństwa i statystyką rozwiązanie osiągnięto dzięki R/exams, który jest pakietem typu open-source dostępnym w darmowym programie statystycznym R. Z racji tego, że kształcenie na odległość odgrywa i będzie odgrywać ważną rolę, poniższy artykuł ma na celu przybliżenie i zachęcenie innych do wykorzystania R/exams w trakcie różnego typu zajęć z przedmiotów matematycznych, ale nie tylko.

1.2. Historia pakietu R/exams

Twórcami pierwszej wersji R/exams byli Grün and Zeileis [1], którzy w 2006 roku na Uniwersytecie Wirtschaftsuniversität we Wiedniu zaprojektowali pakiet, mający sprostać wyzwaniu jakim było przeprowadzenie kursu ze wstępu do statystyki dla ponad 1000 studentów. Przedmiot prowadzony był równolegle przez kilku wykładowców, a celem pakietu R/exams było stworzenie zbioru przykładowych zadań z rozwiązaniami do

samodzielnej nauki oraz generowanie dużej liczby egzaminów w celu zapewnienia indywidualnego testu każdemu studentowi. Materiały te były przygotowywane w formacie PDF.

Z chwilą wzrostu popularności kursów e-learningowych oraz rozwojem systemów zarządzania kształceniem na odległość, a wśród nich platformy Moodle [2], nastąpiła potrzeba stworzenia drugiej wersji pakietu R/exams, posiadającej inne formaty wyjściowe niż tylko PDF. Między innymi format Moodle XML, dzięki któremu stworzone zadania można importować i wykorzystać w quizach na platformie Moodle.

Każde utworzone w R/exams ćwiczenie jest pojedynczym plikiem Sweave (.Rnw) [3] będącym kombinacją kodu R do generowania danych i obliczania rozwiązań oraz języka LaTeX do opisu pytania i rozwiązania z wykorzystaniem zapisu matematycznego. Umożliwia to nie tylko generowanie zróżnicowanych materiałów do samodzielnej nauki i indywidualnych egzaminów, ale także możliwość współpracy nad pulą pytań większej liczby osób prowadzących ten sam przedmiot.

Do chwili obecnej pakiet R/exams został rozszerzony o nową opcję formatowania tekstu pytań i rozwiązań z wykorzystaniem języka Markdown (.Rmd) [4], który jest polecany autorom nieznającym języka LaTeX. Dodatkowo został wprowadzony interfejs `exams2nops()`, dzięki któremu egzaminy napisane w formie tradycyjnej mogą być automatycznie skanowane i oceniane [5].

2. INFORMACJE OGÓLNE O R/EXAMS

Pakiet R/exams oferuje trzy mechanizmy losowań:

- losowy wybór pytań z puli przygotowanych zadań,
- w pytaniach jednokrotnego i wielokrotnego wyboru losowe tasowanie alternatywnych odpowiedzi,
- losowe wybieranie liczby, tekstu, grafiki jako danych do zadania.

Pojedynczy plik z zadaniem składa się z czterech elementów:

- kodu generującego dane,
- tekstu pytania,
- rozwiązania,
- meta-informacji (m.in. etykieta, rodzaj pytania, prawidłowe rozwiązanie).

Generowanie danych oraz podawanie rozwiązania jest opcjonalne, nie musi zostać zawarte w pliku z ćwiczeniem, pozostałe są obowiązkowe.

Do wyboru mamy następujące rodzaje pytań:

- pojedynczy wybór (schoice),
- wielokrotny wybór (mchoice),
- numeryczne (num),
- łańcuch znaków (string),
- kombinacje powyższych (cloze).

Pierwsze dwa są standardowymi pytaniami testowymi umożliwiającymi wybór jednej bądź wielu odpowiedzi.

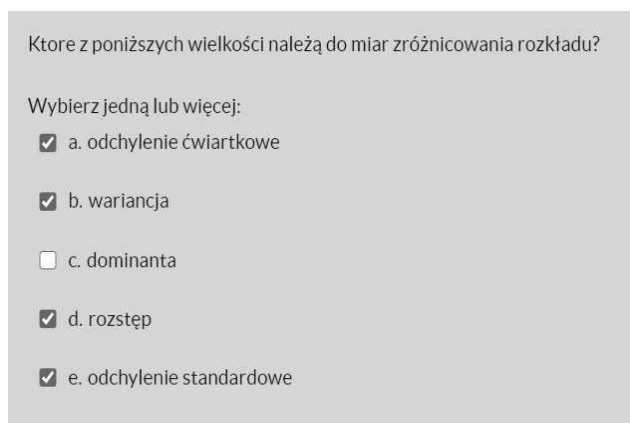
Celem zadań numerycznych jest policzenie i podanie wartości liczbowej. Dopuszcza się w nim podanie pewnego zakresu tolerancji. Z kolei oczekiwaną odpowiedzią do pytań typu łańcuch znaków jest ciąg wyrazów.

W pakiecie R/exams zostały wbudowane przykładowe zadania [6], którymi można się inspirować tworząc bazę swoich własnych pytań z danego przedmiotu. Możemy je wywołać w różnych formatach wyjściowych wykorzystując odpowiednie polecenia, np. `exams2pdf()`, `exams2html()` lub `exams2moodle()`. Plik wyeksportowany w ostatnim formacie następnie jest otwierany na platformie e-nauczania. Jest to sposób na szybkie i efektywne rozszerzenie bazy pytań w swoim kursie.

3. PRZYKŁADOWE MOŻLIWOŚCI R/EXAMS

3.1. Struktura zadania

Na podstawie prostego pod względem budowy pytania wielokrotnego wyboru zostanie zobrazowana struktura kodu napisanego w formacie `.Rnw` oraz `.Rmd`. Składa się ona z trzech elementów: treści pytania, rozwiązania oraz meta-informacji. Ćwiczenie pochodzi z zakresu statystyki opisowej i zawiera 8 odpowiedzi, ale do każdego wariantu pytania jest losowanych i tasowanych 5 z pośród nich, o czym informuje opcja `exshuffle`. Przykładowa wersja widoczna jest na rysunku 1, a wykorzystany kod umieszczony w tablicy 1.



Rys. 1. Pytanie wielokrotnego wyboru stworzone w R/exams i wykorzystane w quizie na platformie e-nauczanie

Drugie niezbyt złożone ćwiczenie, jednak zawierające w swojej budowie dodatkowo czwarty element, jakim jest kod generujący dane, zawiera tablica 2.

Należy zauważyć, że część kodu R odpowiedzialna za losowanie danych została zapisana w pliku `.Rnw` za pośrednictwem `<<>>=` oraz `@`, a w środowisku Markdown w otoczeniu ````{r data generation, echo=FALSE, results="hide"}```` oraz `````. Dodatkowo w obu środowiskach, wykorzystując odpowiednio `\Sexpr{}` oraz ``r...``, umieszczane są dane lub funkcje pochodzące z programu R.

Tablica 1. Kod pytania wielokrotnego wyboru napisany w języku LaTeX i Markdown

LaTeX (.Rnw)	Markdown (.Rmd)
<pre> \begin{question} Które z poniższych wielkości należą do miar zróżnicowania rozkładu? \begin{answerlist} \item odchylenie standardowe \item średnia arytmetyczna \item współczynnik zmienności \item mediana \item dominanta \item rozstęp \item wariancja \item odchylenie ćwiartkowe \end{answerlist} \end{question} \begin{solution} \begin{answerlist} \item prawda \item fałsz \item prawda \item fałsz \item fałsz \item prawda \item prawda \item prawda \item prawda \end{answerlist} \end{solution} \exname{zmiennosc} \extype{mchoice} \exsolution{10100111} \exshuffle{5} </pre>	<pre> Question ===== Które z poniższych wielkości należą do miar różnicowania rozkładu? Answerlist ----- * odchylenie standardowe * średnia arytmetyczna * współczynnik zmienności * mediana * dominanta * rozstęp * wariancja * odchylenie ćwiartkowe Solution ===== Answerlist ----- * prawda * fałsz * prawda * fałsz * fałsz * prawda * prawda * prawda * prawda Meta-information ===== exname: zmiennosc extype: mchoice exsolution: 10100111 exshuffle: 5 </pre>

Tablica 2. Kod pytania numerycznego napisany w języku LaTeX i Markdown zawierający fragment generujący dane

LaTeX (.Rnw)
<pre> <<echo=FALSE, results=hide>>= n <- sample(15:20, 1) w <- sample(c(0:25), n) srednia<-mean(w) @ \begin{question} W \Sexpr{n} osobowej klasie został przeprowadzony sprawdzian z matematyki, z którego można było otrzymać maksymalnie 25 punktów. Uczniowie uzyskali następujące wyniki: \Sexpr{w}. Ile wynosi średnia arytmetyczna punktów uzyskanych przez klasę z tego sprawdzianu? \end{question} \begin{solution} Średnia arytmetyczna wyników ze sprawdzianu wynosi $\Sexpr{round(srednia,2)}$. \end{solution} \extype{num} </pre>

```

\exsolution{\Sexpr{round(srednia,2)}}
\exname{srednia}
\extol{0.01}

```

Markdown (.Rmd)

```

```{r data generation, echo = FALSE,
results = "hide"}
n <- sample(15:20, 1)
w <- sample(c(0:25), n)
srednia<-mean(w)
```

```

Question
=====

W `r` n` osobowej klasie został przeprowadzony sprawdzian z matematyki, z którego można było otrzymać maksymalnie 25 punktów. Uczniowie uzyskali następujące wyniki: `r` w`. Ile wynosi średnia arytmetyczna punktów uzyskanych przez klasę z tego sprawdzianu?

Solution
=====

Średnia arytmetyczna wyników ze sprawdzianu wynosi
`r` round(srednia,2)`.

Meta-information
=====

```

extype: num
exsolution: `r` round(srednia,2)`
exname: srednia
extol: 0.01

```

Tolerancja dla udzielonych odpowiedzi w przytoczonym ćwiczeniu wynosi 0.01 (extol). Przykładowe zadanie wygenerowane przy pomocy powyższego kodu znajduje się na rysunku 2.

W 18 osobowej klasie został przeprowadzony sprawdzian z matematyki, z którego można było otrzymać maksymalnie 25 punktów. Uczniowie uzyskali następujące wyniki: 21, 12, 19, 25, 4, 23, 14, 3, 6, 9, 2, 11, 7, 0, 20, 5, 18, 13. Ile wynosi średnia arytmetyczna punktów uzyskanych przez klasę z tego sprawdzianu?

Odpowiedź:

Rys. 2. Pytanie numeryczne stworzone w R/exams i wykorzystane w quizie na platformie e-nauczanie.

3.2. Zadania typu cloze

Zadania, będące kombinacjami pytań numerycznych, łańcuchowych, jedno- lub wielokrotnego wyboru, są typu cloze. Dają one możliwości stworzenia ciekawych ćwiczeń i pytań egzaminacyjnych. Ich kod jest o wiele bardziej rozbudowany niż zadań przytoczonych w poprzednim podrozdziale, dlatego poniżej ograniczymy się do zaprezentowania jedynie efektów końcowych, jakimi są zadania w formie pdf, html czy pytania w bazie Moodle oraz fragmentów wykorzystanych w ćwiczeniu poleceń.

Zadanie zaprezentowane na rysunku 3 zawiera w sobie dwa zadania numeryczne, co oznacza, że oczekiwaną odpowiedzią są dwie liczby, jakie w tym przypadku stanowią wartość górną i dolną przedziału ufności dla średniej.

Inny wariant tego samego zadania w formacie Moodle widoczne jest na rysunku 4. Liczba elementów próby (liczba klientów) oraz wartości elementów próby (czas obsługi każdego z klientów) generowane są w sposób losowy. Wielkość próby to liczba z przedziału od 50 do 150 ($n <$

$\text{sample}(50:150,1)$). Wartości próby zostały wzięte z rozkładu normalnego z losowymi, pochodzącymi z zadanego przedziału, parametrami średniej oraz odchyleniem standardowym ($y \leftarrow \text{rnorm}(n, \text{runif}(1, 100, 200), \text{runif}(1, 10, 15))$).

Problem

Chcemy oszacować średni czas obsługi przy kasie w pewnym gdańskim markecie. W tym celu dokonano pomiaru czasu obsługi 132 losowo wybranych klientów i otrzymano w rezultacie średnią 134.5 sekund oraz wariancję s_{n-1}^2 równą 113.1.

Wyznacz 95% przedział ufności dla średniej czasu obsługi klientów w badanym sklepie. (Wynik zaokrąglij do trzech miejsc po przecinku, tolerancja dla odp. wynosi ± 0.01)

- Jaka jest dolna granica przedziału ufności?
- Jaka jest górna granica przedziału ufności?

Solution

95% przedział ufności dla średniej μ wynosi:

$$\begin{aligned} & \left[\bar{y} - 1.96 \sqrt{\frac{s_{n-1}^2}{n}}, \bar{y} + 1.96 \sqrt{\frac{s_{n-1}^2}{n}} \right] \\ & = \left[134.5 - 1.96 \sqrt{\frac{113.1}{132}}, 134.5 + 1.96 \sqrt{\frac{113.1}{132}} \right] \\ & = [132.686, 136.314]. \end{aligned}$$

- Dolna granica przedziału ufności to 132.686.
- Górna granica przedziału ufności to 136.314.

Rys. 3. Przykładowe pytanie wraz z rozwiązaniem w formacie wyjściowym pdf

Chcemy oszacować średni czas obsługi przy kasie w pewnym gdańskim markecie. W tym celu dokonano pomiaru czasu obsługi 50 losowo wybranych klientów i otrzymano w rezultacie średnią 122.1 sekund oraz wariancję s_{n-1}^2 równą 137.2.

Wyznacz 95% przedział ufności dla średniej czasu obsługi klientów w badanym sklepie. (Wynik zaokrąglij do trzech miejsc po przecinku, tolerancja dla odp. wynosi ± 0.01)

- Jaka jest dolna granica przedziału ufności?
- Jaka jest górna granica przedziału ufności?

Rys. 4. Pytanie typu cloze (podwójne numeryczne) stworzone w R/exams i wykorzystane w quizie na platformie e-nauczanie

Wartość średnia oraz wariancja podane w zadaniu są uzależnione od wielkości i wartości wylosowanej próby i zostały wyliczone z wykorzystaniem funkcji R.

Inne ciekawe zadanie łączące w sobie pytania jednokrotnego wyboru oraz pytania numeryczne, jak również dające możliwość weryfikacji umiejętności studentów w obsłudze programu statystycznego wykorzystywanego na zajęciach laboratoryjnych, zostało zaproponowane na rysunku 5. Początkowo zostaje utworzony plik typu .csv z, podobnie jak poprzednio, losową ilością i wielkością danych. W poszczególnych wariantach zadania zmiana ulega również zakres tolerancji.

Plik pudełko.csv zawiera 42 obserwacji zmiennej x. Na ich podstawie narysuj wykres pudełkowy i odpowiedz na poniższe pytania (Tolerancja dla odpowiedzi, w których należy wpisać wartości wynosi ± 0.2):

Z wykresu pudełkowego można odczytać rozstęp danych:

Rozkład badanej zmiennej jest:

Wykres pudełkowy pokazuje wartości odstające:

Jedna czwarta obserwacji ma wartości nie większe niż:

Jedna czwarta obserwacji ma wartości nie mniejsze niż:

Półowa obserwacji ma wartości nie mniejsze niż:

Rys. 5. Pytanie typu cloze (jednokrotny wybór, numeryczne)

Student w pierwszym kroku otwiera wygenerowane dane i rysuje wykres na podstawie, którego następnie udziela odpowiedzi na zadane pytania. Rysunek 6 prezentuje inny wariant tego samego ćwiczenia w formacie html.

1. Question

Plik `pudelko.csv` zawiera 49 obserwacji zmiennej x . Na ich podstawie narysuj wykres pudełkowy i odpowiedz na poniższe pytania (*Tolerancja dla odpowiedzi, w których należy wpisać wartości wynosi ± 0.8*):

Z wykresu pudełkowego można odczytać rozstęp danych: tak / nie

Rozkład badanej zmiennej jest: symetryczny / prawostronnie asymetryczny / lewostronnie asymetryczny

Wykres pudełkowy pokazuje wartości odstające: tak / nie

Jedna czwarta obserwacji ma wartości nie większe niż: tak / nie

Jedna czwarta obserwacji ma wartości nie mniejsze niż: tak / nie

Półowa obserwacji ma wartości nie mniejsze niż: tak / nie

Solution

a. Fałsz. / Prawda.
 b. Prawda. / Fałsz. / Fałsz.
 c. Fałsz. / Prawda.
 d. 2.55.
 e. 13.1.
 f. 10.4.

Rys. 6. Przykładowe pytanie wraz z rozwiązaniem w formacie wyjściowym html

3.3. Ułożenie treści zaliczenia

Układając egzamin należy wskazać pliki z zadaniami, które mają zostać zawarte w treści: `zaliczenie<-c("zmiennosc.Rmd", "srednia.Rmd", "przedzial.Rmd", "pudelko.Rmd")`, a następnie wyeksportować w formacie Moodle XML, definiując dowolną liczbę ich replikacji `exams2moodle(zaliczenie, n = 10, name = "Egzamin")`. Utworzona w ten sposób duża ilość pytań, po zaimportowaniu w kursie na e-nauczanie, jest dzielona na katalogi według różnych wariantów tego samego zadania. Następnie, wykorzystując możliwości Moodle, dane pytanie

w quizie może być losowane z każdego katalogu, co zapewnia niepowtarzalność egzaminów.

4. WNIOSKI KOŃCOWE

Ręczne ułożenie i sprawdzenie wielu wariantów kolokwium czy egzaminów jest pracochłonne, a w ilości zapewniającej indywidualny test dla każdego studenta, wręcz niemożliwe. R/exams daje szerokie możliwości ułożenia nie tylko zadań na zaliczenie, ale również przygotowania ćwiczeń z rozwiązaniami do samodzielnej nauki. Dzięki pakietowi jest możliwe stworzenie różnego typu zadań, ale przede wszystkim kilkunastu wariantów tego samego typu, aby uniknąć uczenia się na pamięć oraz ściągania w trakcie egzaminów. Artykuł pokazuje jedynie kilka przykładowych zadań stworzonych w R/exams, aby przybliżyć możliwości jakie możemy dzięki niemu osiągnąć. Dodatkowo fakt, że pakiet pozwala na produkcję zadań w formacie Moodle XML, czyni go atrakcyjnym pod względem użytkowania dla każdego nauczyciela, szczególnie w momencie, gdy zdalne nauczanie i przeprowadzanie zaliczeń stało się codziennością.

5. BIBLIOGRAFIA

1. Grün B., Zeileis A.: Automatic Generation of Exams in R, Journal of Statistical Software, Nr 29 (10), 2009, s. 1-14. <http://www.jstatsoft.org/v29/i10/>.
2. Dougiamas M, et al.: Moodle, Version 2.6. 2014. <http://moodle.org/>.
3. Leisch F.: Sweave: Dynamic Generation of Statistical Reports Using Literate Data Analysis, In: Härdle W., Rönz B. (eds) Compstat. Physica, Heidelberg, 2002, s.575-580 https://doi.org/10.1007/978-3-642-574894_89.
4. Fireball D.: Markdown Syntax Documentation, <https://daringfireball.net/projects/markdown/syntax>.
5. Zeileis A., Umlauf N., Leisch F.: Flexible Generation of E-Learning Exams in R: Moodle Quizzes, OLAT Assessments, and Beyond, 2020 <https://cran.r-project.org/web/packages/exams/vignettes/exams2.pdf>.
6. <http://www.r-exams.org/templates/>.

R/EXAMS AS SUPPORT IN VERIFICATION OF STUDENTS' KNOWLEDGE

The usefulness of e-learning resources, such as online tests and quizzes, can be seen in many situations: a large number of students attending a given course, consolidating, summarizing and verifying of knowledge or distance learning. The aim of the article is to present the R/exams package available in the R statistical program as a support for creating random and dynamic questions and their use on the Moodle platform. Thanks to the package, it is possible to create various types of tasks, but above all, a dozen variants of the same type in order to avoid rote learning and cheating during exams. Due to the fact that distance education plays and will play an important role, the article aims to introduce and encourage others to use R/exams during various types of mathematics courses, but not only. In addition, the varied output forms such as PDF, HTML or Moodle XML make it very attractive in terms of use. In this article various types of questions are discussed using the example of statistical tasks.

Keywords: exam, test, "exams" package, Moodle.