

Personalized exams in Probability and Statistics

Gallusz Abaliget, Dániel Kehl

University of Pécs

09-09-2017



PÉCSIKÖZGÁZ
utoljára a karrier indít

Statistics related courses

Overview

- ▶ BA in Hungarian
 - ▶ Valószínűségszámítás és statisztika (Probability and Statistics)
 - ▶ 400/200 students, 1 lecture 7-8/3-4 computer lab sessions
 - ▶ Statisztikai modellezés (Statistical Modeling)
 - ▶ 100/250 students, 1 lecture 2/4-5 computer lab sessions
 - ▶ *Makrogazdasági adatok stat. el. (Applied Economic Statistics)*
 - ▶ 30-40 students, 1 lecture
 - ▶ *Statisztikai modellezés R-ben (Statistical Modeling in R)*
 - ▶ 20-30 students, 1 computer lab session
- ▶ BA in English
 - ▶ Probability and Statistics
 - ▶ 60-70 students, 1 lecture 2 computer lab sessions (autumn)
 - ▶ Business Statistics
 - ▶ 60-70 students, 1 lecture 2 computer lab sessions (spring)

Main challenges

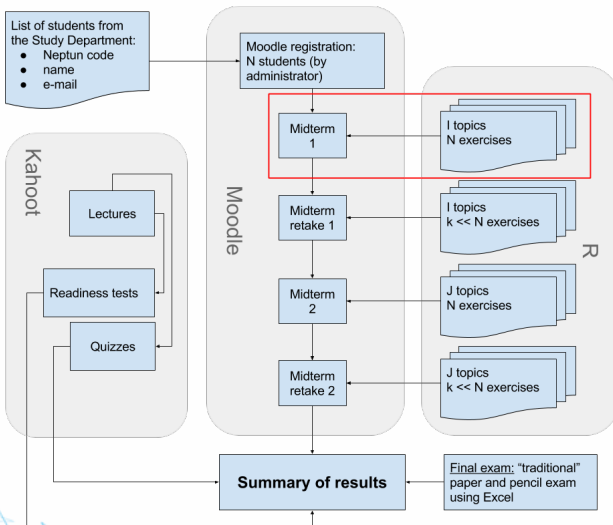
in case of relatively big classes

- ▶ high number of student especially on the Hungarian Programme
 - ▶ high number of exams, takes a lot of time grading, creating new problems, datasets etc.
 - ▶ after midterms students want to check their exams, solutions etc.
- ▶ desire to „force” students to continuously follow the course material throughout the semester
 - ▶ in-class short, 5 minute quizzes every lab session
 - ▶ two Excel-based midterms
- ▶ exams at the computer makes cheating easier

One possible solution is personalized exams with Moodle and R.

Personalized midterms

Assessment structure of the semester



Personalized midterms

Storing exercise text and data in spreadsheets

The example of estimation/hypothesis testing of the population mean

text	text2	mu	sigma	round
Egy üzleti döntés előkészítéséhez meg kell becsülnünk a termékünk ellenőrzéséhez szükséges átlagos időt (percben). Ehhez az alábbi nnnn elemű mintát választottuk.	Egy üzleti döntés előkészítéséhez meg kell vizsgálnunk a termékünk ellenőrzéséhez szükséges átlagos időt (percben). Ehhez az alábbi nnnn elemű mintát választottuk. A termelési igazgató állítása szerint az ellenőrzéshez szükséges átlagos idő mumumu	100	15	1
Egy nagyvállalatnál éves TeljesítményÉrtékelési Rendszer (TÉR) működik. A kitöltéshez szükséges átlagos időt kívánjuk megbecsülni. Véletlenszerűen kiválasztott nnnn munkavállaló esetén mértük ezt az időt.	Egy nagyvállalatnál éves TeljesítményÉrtékelési Rendszer (TÉR) működik. A kitöltéshez szükséges átlagos időt kívánjuk vizsgálni véletlenszerűen kiválasztott nnnn munkavállaló segítségével. A TÉR-t szállító külső vállalat állítása szerint az	25	5	1
A fogyasztói árindex számításához szükségünk van az egy kilogrammos fehér kenyér magyarországi átlagára. Ennek érdekében az alábbi reprezentatív árszazeírásokat végeztük.	Az egyik politikai párt szerint az egy kilogrammos fehér kenyér magyarországi átlagára mumumu forint. Ennek ellenőrzésére az alábbi reprezentatív árszazeírásokat végeztük.	250	20	0
Egy vállalatnak készített marketing	Egy vállalatnak készített marketing			

Personalized midterms

The R-code – generating data and solution

```
1 >> <<echo=FALSE, results=hide>>=
2 ## DATA GENERATION
3 szovegek <- read.csv2(file.path(mywd, "exercises", "prob_stat", "estimation", "mean_est
  hip.csv"), stringsAsFactors = FALSE)
4 szovegek <- szovegek[sample(1:nrow(szovegek), 1), ]
5 signLevel <- sample(c(.1, .05, .01), 1)
6
7 n <- sample(x = 6:8, size = 1)*5
8 smp1 <- round(rnorm(n, szovegek$mu, szovegek$sigma), szovegek$round)
9
10 data <- matrix(smp1, ncol = 5)
11
12 dataDisp <- xtable(data, digits = szovegek$round)
13
14 ## CALCULATIONS
15 m <- mean(smp1)
16 se <- sd(smp1) / sqrt(n)
17
18 mu0 <- round(m - runif(1, -4, 4) * se, szovegek$round)
19
20 temp <- (m-mu0)/se
21 tkrit <- -qt((signLevel) / 2, n - 1)
22
23 pertek <- 2*(1-pt(abs(temp), n - 1))
```

Personalized midterms

The R-code – generating questions, setting tolerances

```
25 ## QUESTION/ANSWER GENERATION
26 questions <- character(4)
27 solutions <- logical(4)
28 tolerances <- rep(0.0001, 4)
29
30 questions[1] <- "Mekkora a hipotézisellenőrzés során használt sztenderd hiba nagysága?"
31 solutions[1] <- se
32
33 questions[2] <- paste0("Adja meg a ", 100*signLevel, "\\%-os szignifikancia szinthez
34 tartozó kétoldalú alternatív hipotézishez tartozó kritikus érték abszolút értékét!")
35 solutions[2] <- tkrit
36
36 questions[3] <- "Adja meg a próbafüggvény empirikus értékét!"
37 solutions[3] <- temp
38
39 if(runif(1) > 0.5) {
40   questions[4] <- "Adja meg a kétoldalú alternatív hipotézishez tartozó p-értéket!"
41   solutions[4] <- pertek
42 } else {
43   questions[4] <- "Adja meg az egyoldalú kisebb alternatív hipotézishez tartozó p
44   -értéket!"
45   solutions[4] <- pt(temp, n - 1)
46 }
```

Personalized midterms

Output in Moodle

Egy nagyvállalatnál éves TeljesítményÉrtékelési Rendszer (TÉR) működik. A kitöltéshez szükséges átlagos időt kívánjuk megbecsülni. Véletlenszerűen kiválasztott 30 munkavállaló esetén mértük ezt az időt.

21,7 30,5 22,4 28,7 20,3
24,7 17,0 20,8 29,9 27,5
28,4 15,7 27,8 23,9 33,0
28,0 21,4 21,1 20,4 15,6
26,7 27,9 19,2 27,2 22,0
22,8 25,5 32,2 24,6 25,6

- Adja meg a mintabeli szórást!
- Mekkora a becslés hibahatára 90%-os megbízhatóság mellett?
- Adja meg a 90%-os megbízhatósági szintű becslés konfidencia intervallumának alsó határát!

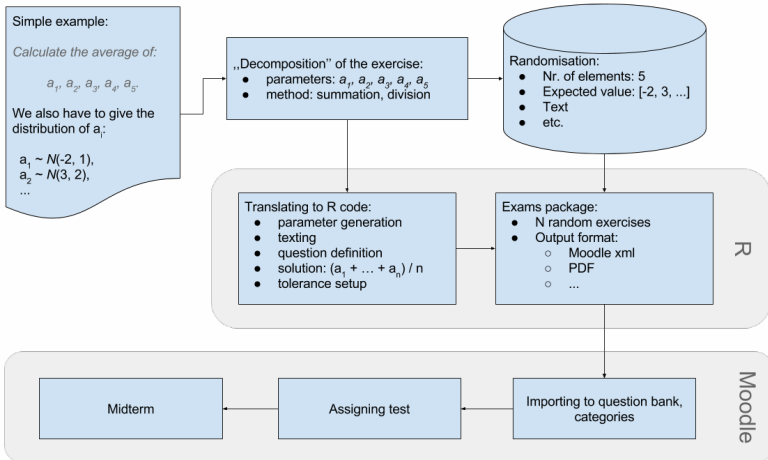
Egy nagyvállalatnál éves TeljesítményÉrtékelési Rendszer (TÉR) működik. A kitöltéshez szükséges átlagos időt kívánjuk megbecsülni. Véletlenszerűen kiválasztott 35 munkavállaló esetén mértük ezt az időt.

23,8 28,6 24,2 19,9 31,3
24,8 23,3 14,4 27,0 23,9
20,3 23,8 24,6 21,2 17,2
23,4 33,4 24,3 21,2 26,1
30,6 28,0 28,3 34,4 23,8
31,2 24,5 23,6 26,3 30,4
24,3 26,2 25,0 22,6 24,8

- Adja meg a mintaátlagot!
- Mekkora a becslés sztenderd hibájának nagysága?

Personalized midterms

The general idea – workflow



Personalized midterms

The R-code – generating a midterm

```
1 source("functions.R")
2
3 mywd <- getwd()
4 folder <- "exercises/prob_stat/estimation/"
5 exerc1617osz1 <- c("binom.Rnw", "hipgeom.Rnw", "poisson.Rnw", "norm.Rnw", "2valt.Rnw")
6 exerc1617osz2 <- c("mean_est.Rnw", "prop_est.Rnw", "prop_hip.Rnw", "f_egyez.Rnw")
7 exerc1617tavasz1 <- c("leiro_stat.Rnw", "binom.Rnw", "hipgeom.Rnw", "poisson.Rnw")
8 exerc1617tavasz1pot <- c("leiro_stat.Rnw", "binom.Rnw", "hipgeom.Rnw", "poisson.Rnw")
9 exerc1617tavasz2 <- c("mean_est.Rnw", "mean_hip.Rnw", "paros.Rnw", "prop_est.Rnw", "prop_
10
11 exerc <- c("mean_est.Rnw", "prop_hip.Rnw", "f_egyez.Rnw")
12
13 myexam <- paste0(folder, exerc)
14
15 exams2pdf(myexam, n = 20, name = c(paste0(c(sub(".Rnw","",unique(exerc)),"exam"), collapse=""),
16                                     paste0(c(sub(".Rnw","",unique(exerc)),"solution"), collapse="")),
17         encoding = "UTF-8",
18         edir = "exercises",
19         dir = "output",
20         template = c("templates/exam.tex", "templates/solution.tex"),
21         header = list(
22           Date = "2017-09-09",
23           ID = function(i) formatC(i, width = 5, flag = "0")
24         )
25     )
26
27 exams2moodle(myexam, n = 50, name = c("szeged"),
28             encoding = "UTF-8",
29             edir = "exercises",
30             dir = "output")
```

As a result

our students face

- ▶ very similar question types (let's say one-sample t-test, ANOVA, estimation and linear regression)

BUT

- ▶ different „stories”
- ▶ different questions (give the empirical value vs. give the critical value vs. give the p-value etc.)
- ▶ different datasets
- ▶ different numeric solutions
- ▶ immediate feedback and results
- ▶ possibility to „flag” questions in Moodle

Summary of our experiences

- ▶ results are fairly similar in comparison to previous years
- ▶ students do not complain about it, like the quick response
- ▶ preparing a midterm takes longer (see R code)
- ▶ setting up a question bank is an initial investment
- ▶ going through flagged questions is fairly quick
- ▶ saving a lot of time with automatic grading
- ▶ cheating seems to be harder

Useful links and materials

- ▶ <https://moodle.org/>
- ▶ R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- ▶ Achim Zeileis, Nikolaus Umlauf, Friedrich Leisch (2014). Flexible Generation of E-Learning Exams in R: Moodle Quizzes, OLAT Assessments, and Beyond. *Journal of Statistical Software* 58(1), 1-36. doi:10.18637/jss.v058.i01
- ▶ <https://cran.r-project.org/web/packages/exams/vignettes/exams.pdf>
- ▶ <https://cran.r-project.org/web/packages/exams/exams.pdf>
- ▶ exams_skeleton function
- ▶ abaligetig@ktk.pte.hu, kehld@ktk.pte.hu