

Informe Final: Projecte 2020PID-UB/017

Millorar l'aprenentatge de Disseny d'Experiments per la incorporació de l'avaluació continuada mitjançant qüestionaris personalitzats

Responsable: Subirana, Isaac

Participants: Calvo, Miquel, Pizarro, Víctor; Pijuan, Pau; Tort-Martorell, Xavier; Sánchez, José Antonio; Camps, Oriol.

Departament: Genètica, Microbiologia i Estadística, Secció Estadística.

Ensenyament: Grau interuniversitari UB-UPC d'Estadística

Assignatura: Disseny d'Experiments

Data d'inici i finalització: 1 d'octubre de 2020 a 1 de juny de 2022

Grup/s d'Innovació Docent implicat/s en l'actuació: Grup d'Innovació Docent Consolidat Statmedia.

Responsable: Vegas, Esteban

Resum

En aquest projecte s'ha explorat i utilitzat les eines del paquet d'R exams per tal de generar activitats personalitzades per als alumnes de l'assignatura de Disseny d'Experiments del grau interuniversitari UB-UPC d'Estadística. El paquet exams permet crear qüestionaris que incorporin preguntes de tipus test, numèriques o una barreja d'ambdues. Mitjançant codi R es poden generar aleatòriament diferents conjunts de dades i possibles respostes de tal manera i així els qüestionaris siguin únics per a cada alumne. Una altra avantatge que ofereix el paquet exams és que en integrar-se al Moodle un cop finalitzat el qüestionari la nota queda integrada automàticament a la base de dades de qualificacions. A més, l'alumne veu immediatament la nota i quins errors ha comès i les solucions de cada resposta. Tot i que els professors no teníem experiència prèvia amb el paquet exams, no ha sigut difícil confeccionar les preguntes ni tampoc integrar-les a la plataforma del campus virtual a través del Moodle. Al contrari, aquest procés ha resultat senzill i no hem trobat cap dificultat tècnica. L'alumnat ha respost de forma positiva, i han trobat gratificant el fet de saber la nota i els errors que havien comès en acabar el qüestionari. Per aquest motiu, creiem que l'ús del paquet exams com a part de la docència de l'assignatura és molt positiva i planegem continuar usant-lo tot explorant noves funcionalitats que ofereixi aquesta eina. Per últim, pensem que la utilització d'exams es podria estendre també a altres assignatures de anàlisi de dades i d'altres graus, com de fet s'està realitzant a l'assignatura de Disseny d'Experiments i Anàlisi de Dades (DEAD) del grau de Biomedicina.

Paraules clau: Design of Experiments, Moodle quizzes, individualized activities, automatic assessment, R, exams.

1. Introducció

L'assignatura de Disseny Experimental és una assignatura obligatòria de tercer curs del grau interuniversitari UB-UPC d'Estadística. Aquesta assignatura és impartida a parts iguals en hores de docència i contingut pel departament d'Enginyeria Industrial de la UPC i per al departament de Microbiologia, Genètica i Estadística de la UB. Part del contingut de l'assignatura és compartida amb l'assignatura de Disseny d'Experiments i Anàlisi de Dades (DEAD), dels graus de Biologia o Biomedicina, que també són impartides pel departament de Microbiologia, Genètica i Estadística, amb una dilatada i àmplia experiència en aquesta assignatura.

Disseny d'Experiments és una assignatura eminentment pràctica en què la teoria es va explicant mitjançant exemples que es van discutint i resolent a classe. Una de les habilitats més importants que ha d'adquirir l'alumne és el maneig de les funcions i eines d'R per a resoldre i analitzar les dades d'un disseny experimental.

L'ús d'eines per a generar exemples individualitzats es converteix en una possibilitat molt interessant tant perquè l'alumne pugui reforçar el coneixement que va adquirint durant el curs i per a poder avaluar-los o fins i tot que ells s'autoavaluïn. En aquest sentit el departament de Microbiologia, Genètica i Estadística va confeccionar l'eina Statmedia que es va fer servir durant molts cursos, però que actualment està en desús per a motius tècnics.

Anteriorment, l'alumnat realitzava exàmens parcials i treballs que eren entregats a mà. Les preguntes dels exàmens parcials i dels treballs sovint eren de resposta a elaborar, la qual cosa demandava una quantitat d'hores considerables al professorat, donat el nombre d'alumnes matriculats: els últims anys el nombre de matriculats a l'assignatura ha sigut d'entre 50 i 70 alumnes, mentre que les classes es repartien entre dos professors fent la meitat de les classes cadascun d'ells. Tot això afegit a la inevitable component subjectiva que comporta possibles errors a l'hora de corregir.

Per tal motiu es van explorar alternatives com el paquet exams d'R (<https://www.r-exams.org/>) que permet confeccionar de forma automatitzada exercicis individualitzats tot aprofitant la potencialitat i versatilitat del programari R i que fàcilment es pot incorporar al Moodle (campus virtual de l'assignatura). R és un programari gratuït, de codi lliure que es pot usar en qualsevol sistema operatiu i que és el més usat en diferents àmbits de la recerca i també en la docència.

La opció del paquet exams s'ha incorporat als darrers cursos en la part d'avaluació continuada, concretament als exàmens parcials i qüestionaris de prova d'auto-avaluació. No només ha restat molta càrrega de treball i reduït problemes en la correcció per part del professorat, sinó que a més el retorn de l'alumnat ha sigut molt satisfactori. El punt que valoren més els alumnes com a positiu és la immediatesa en saber la nota i també en saber on s'han equivocat i el motiu. A més, en cas de possible dubte o disconformitat en alguna correcció, és possible rectificar o modificar la puntuació concreta d'una pregunta en cas que sigui oportú, de forma molt fàcil.

2. Actuació d'innovació docent

2.1 Context d'aplicació

Els alumnes que cursen l'assignatura de Disseny d'Experiments tenen un grau de coneixement molt heterogeni, tant pel que fa a aspectes teòrics, com més pràctics o d'eines de software. En general, en finalitzar el curs, la majoria assoleixen el nivell necessari per a aprovar l'assignatura i/o tenir una qualificació de notable o excel·lent. Els recursos que s'ofereixen, tant d'espai com de material informàtic és satisfactori i permet desenvolupar les classes sense cap entrebanc. D'altra banda, alguns alumnes opten per dur el seu propi ordinador, cosa que creiem que és bona ja que evita haver de re-instal·lar cada vegada alguns paquets o programaris que es poden esborrar en iniciar sessions diferents si s'usen els equips que ofereix la facultat.

En general, els alumnes que assisteixen assíduament a classe aprenen tot o gran part del contingut del curs, mostrant en general interès per a l'assignatura. I valoren sobretot el caire pràctic que es dona a l'assignatura i l'ús d'R en la resolució dels problemes.

2.2 Objectius

1. *Tecnològic*: explorar tota la potencialitat del paquet exams per tal de confeccionar qüestionaris i integrar-los al campus virtual.
2. *Activitats*: les activitats planificades amb exams consisteixen en dos exàmens parcials i qüestionaris que cobreixen els diferents temes de l'assignatura.
3. *Aules*: els exàmens parcials es faran en horari docent i en aules equipades amb ordinadors, mentre que els qüestionaris de pràctica o d'auto-avaluació es farà fora d'hores docents.

2.3 Metodologia i desenvolupament de l'actuació

Els professors de l'assignatura disposàvem de coneixements tècnics previs de R i R-markdown. Un d'ells va assistir a un seminari sobre R+exams al congrés d'R (use-R 2019). A més, tots els qüestionaris creats amb R+exams han sigut confeccionats pels mateixos professors encarregats de l'assignatura.

Primerament, s'han reproduït alguns exemples disponibles al material d'ajuda i manuals del paquet exams (<https://www.r-exams.org/tutorials/>) per tal de familiaritzar-nos amb les funcions d'aquesta eina. En segon lloc, s'han adaptat alguns exemples de la llista de problemes de l'assignatura. Posteriorment s'han generat alguns qüestionaris específics per a l'assignatura i per a poder-los incorporar al Moodle. Finalment s'han creat una bateria de preguntes. Algunes d'elles serviran pels qüestionaris de pràctica (d'auto-avaluació) per als estudiants i les altres per als exàmens parcials. La creació de les preguntes del banc de proves, ha comportat el 80% o 90% del volum de treball, mentre que la resta ha sigut la confecció pròpiament de cada qüestionari al Moodle tot escollint quines preguntes i la seva puntuació, i la validació per tal de comprovar que tot està correctament implementat. Cadascun dels dos professors encarregats de

l'assignatura ha creat preguntes per als temes que imparteix. D'aquesta manera es reparteix el treball de forma equitativa.

2.4 Recursos i suports emprats

No ha sigut necessari cap recurs addicional ni de material, ni de personal. El professorat encarregat de l'assignatura ja tenia uns coneixements de base suficients per tal de dur a terme el projecte. A més, el paquet exams, com tots els paquets d'R és una eina gratuïta i que no requereix de llicència per tal d'instal·lar-la i fer-la servir. Per altra banda, les versions de Moodle en què es sustenta el campus virtual de l'assignatura és compatible amb les funcions d'R exams i no hi ha hagut cap problema tècnic perquè els alumnes responguessin als qüestionaris.

3. Resultats i conclusions

3.1 Estat del projecte

1. Tecnològic: l'objectiu del projecte fou explorar les possibilitats que podia oferir el paquet d'R exams i com incorporar els qüestionaris generats al Moodle del campus virtual. El primer curs en què es començar a incorporar els qüestionaris a l'assignatura fou el 19-20. Des d'aleshores, diversos professors sense experiència prèvia amb el paquet exams han confeccionat diferents qüestionaris sense dificultat, i l'experiència ha sigut molt bona. La integració de les preguntes al campus virtual s'ha fet sense problemes en tots els cursos. Des del punt de vista de l'alumnat, gairebé no hi ha hagut reclamacions en quant a les notes i correccions de les preguntes, i quan ha calgut fer alguna modificació ha sigut molt fàcil. La gran majoria del comentaris dels alumnes han sigut positius, destacant el fet que puguin saber la nota i en què s'han equivocat tot just acabar els qüestionaris. Tampoc s'ha detectat cap problema a l'hora d'iniciar i acabar el qüestionari ni en el registre de les notes guardades al campus.
2. Activitats: des del curs 19-20, s'han realitzat dos parcials amb exams, el primer aproximadament a meitat del curs i el segon a les darreres classes del curs. A més, s'han creat dos o tres qüestionaris també amb exams que els alumnes podien fer fora d'horari docent algunes setmanes abans de cada parcial.
3. Aules: donat el nombre d'alumnes matriculats i que optaven per a l'avaluació continuada, i per tal de garantir l'accés a ordinadors per a tots els alumnes, cadascun dels dos torns es va haver de dividir en dos en algunes ocasions. El format i caràcter individualitzat dels qüestionaris creats amb exams ha permès resoldre perfectament aquest hàndicap.

A continuació es mostra un exemple de qüestionari, en concret del primer parcial, on es pot veure l'enllaç d'accés a les preguntes (figura 1), les notes per a cada alumne i pregunta (figura 2), i les respostes en detall per a una determinada pregunta i alumne (figures 3a i 3b). A la figura 3a apareix un enllaç per a poder re-avaluar la pregunta i afegir un comentari perquè sigui visualitzat per l'alumne es cas de revisió (figura 3b).

Primer Parcial

→ 👁 ✕ 🔊

Utilitzeu aquesta àrea per descriure sobre què tracta aquest tema (amb text, imatges, àudio i vídeo).

 Edita la secció

QÜESTIONARI ✔

 Primer torn

40 de 69 intentades Data límit 27 d'octubre 2022

→ ✎ ⋮

QÜESTIONARI ✔

 Segon torn

24 de 69 intentades Data límit 27 d'octubre 2022

→ ✎ ⋮

+
Crea activitats d'aprenentatge

 Deixeu anar fitxers per adjuntar o navegueu

Figura 1. Qüestionari de la primera prova parcial del curs 22-23.

Cognoms /		Començat			Temps		P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7
Nom	Número ID	Estat	el	Completat	emprat	Qualificació/10,00	/1,25	/1,25	/1,25	/1,25	/0,31	/0,31	/0,31
		Acabat	27 d'octubre 2022 15:00	27 d'octubre 2022 16:43	1 hora 42 minuts	7,08	✔ 1,25	✔ 1,25	✔ 1,25	✔ 1,25	✘ -0,10	✔ 0,31	✔ 0,31
		Acabat	27 d'octubre 2022 15:00	27 d'octubre 2022 16:17	1 hora 17 minuts	8,85	✔ 1,25	✔ 1,25	☑ 0,94	✔ 1,25	✘ -0,10	✔ 0,31	✔ 0,31
		Acabat	27 d'octubre 2022 15:00	27 d'octubre 2022 16:16	1 hora 16 minuts	9,58	✔ 1,25	✔ 1,25	✔ 1,25	✔ 1,25	✘ -0,10	✔ 0,31	✔ 0,31
		Acabat	27 d'octubre 2022 15:00	27 d'octubre 2022 16:20	1 hora 20 minuts	8,65	✔ 1,25	☑ 1,04	☑ 0,94	✔ 1,25	✔ 0,31	✔ 0,31	✔ 0,31
		Acabat	27 d'octubre 2022 15:00	27 d'octubre 2022 15:48	47 minuts 52 segons	8,43	☑ 1,07	✔ 1,25	✔ 1,25	☑ 1,11	✘ -0,10	✘ -0,10	✔ 0,31
		Acabat	27 d'octubre 2022 15:00	27 d'octubre 2022 16:55	1 hora 54 minuts	9,17	✔ 1,25	✔ 1,25	✔ 1,25	✔ 1,25	✔ 0,31	✔ 0,31	✔ 0,31

Figura 2. Mostra d'algunes notes del qüestionari del primer parcial.

Un grup d'innovació docent va decidir provar 3 mètodes docents per una assignatura. Es van crear 3 grups de classe, i a cada grup de classe es feia servir un dels 3 mètodes docents. Els estudiants es van assignar aleatòriament als grups. Al final, tots els estudiants van fer el mateix examen. L'objectiu és veure si hi ha diferències en promig entre metodologies docents. El nivell de significació és $\alpha = 0.05$.

Aquí teniu les notes finals dels estudiants dels 3 grups:

```
grup1 = c(6.4, 6.6, 7.0, 7.2, 6.5, 7.1, 7.3)
grup2 = c(7.4, 7.3, 7.5, 7.3, 6.5, 6.7, 7.4, 8.3)
grup3 = c(7.1, 7.5, 6.6, 6.8, 6.4, 7.0, 7.6)
df = data.frame(notes = c(grup1, grup2, grup3),
                grup = factor(rep(1:3, c(7,8,7))))
```

Responeu les preguntes següents, **fent servir la coma per separar decimals**:

- a. Quina és la mida mostral de l'experiment?
- b. Quan val el p-valor de la taula ANOVA?
- c. És un disseny equilibrat o desequilibrat?
- d. Hi ha evidència estadística que hi hagi diferències en promig entre les metodologies docents?
- e. Detectem diferències entre les metodologies 1 i 2?
- f. Detectem diferències entre les metodologies 1 i 3?
- g. Detectem diferències entre les metodologies 2 i 3?

- a. La mida mostral és: ✓
- b. El p-valor val: ✓
- c. Desequilibrat ✓
- d. No hi ha diferències entre metodologies ✓
- e. No detectem diferències entre metodologies 1 i 2 ✓
- f. No detectem diferències entre metodologies 1 i 3 ✓
- g. No detectem diferències entre metodologies 2 i 3 ✓

[Fes un comentari o torna a puntuar](#)

Figura 3a. Enunciat i respostes d'una pregunta del qüestionari.

Comentari

Rich text editor toolbar with icons for bold, italic, underline, link, unlink, image, video, table, list, indent, outdent, undo, redo, and other editing functions.

Puntuat

sobre 4,00

Figura 3b. Finestra on es pot modificar la nota d'una pregunta del qüestionari.

3.2 Indicadors d'avaluació del projecte i difusió dels resultats

Per a avaluar l'eficàcia del projecte s'ha tingut en compte diferents indicadors objectivables i d'altres més subjectius o de percepció. Entre els primers s'ha intentat donar una comparativa entre els anys en què s'ha implementat els qüestionaris amb exams i els anys anteriors tot comparant (1) el grau de participació calculat com el percentatge d'alumnes que han triat avaluació continuada i (2) la nota final de l'assignatura entre els alumnes d'avaluació continuada. Com a indicadors més subjectius s'han tingut en compte els comentaris de forma més qualitativa.

Els resultats sobre el percentatge d'alumnes que van optar per a l'avaluació continuada i la nota mitjana de l'assignatura entre aquests alumnes es recullen a la taula 1, tot distingint entre els tres cursos anteriors a la implementació dels qüestionaris amb exams i els tres anys posteriors.

	Abans	Després	p-valor
Mitjana	6,17	6,68	0,009
%Aprovats	85,4	86,9	0,742
%Notables	29,7	49,7	<0,001
%Excel·lents	5,1	6,9	0,628
%Continuada	84	94,8	0,002

Taula 1. Notes obtingudes els tres cursos anteriors a la implementació d'exams (columna "Abans") i els tres cursos en què s'ha implementat l'exams (columna "després"). La nota mitjana i el % d'aprovat, notables i excel·lents s'han calculat entre els alumnes que han optat per l'avaluació continuada.

Com es pot veure a la taula 1, la proporció d'alumnes que han triat fer avaluació continuada ha crescut en els darrers tres cursos que es porta implementant-se l'exams comparat amb els tres cursos anteriors. També es pot observar com la nota mitjana ha crescut així com la proporció de notables i excel·lents entre els que han optat per l'avaluació continuada.

D'entre els comentaris efectuats pels alumnes, cal destacar com a negatiu, el fet que la nota de en algunes ocasions no estava d'acord amb el seu coneixement, ja que, per la naturalesa d'aquest tipus de qüestionaris automatitzats no es pot avaluar els passos intermedis ni el raonament en la resolució dels problemes, sinó només la resposta final. I com a comentaris positius, la immediatesa en la correcció. Per altra banda, un altre aspecte molt important a destacar, és la dedicació en quant al professorat: la càrrega de feina és quasi monopolitzada per la confecció de les preguntes i disseny dels qüestionari, deixant com a residual les hores de correcció.

3.3 Algunes notes tècniques

R+exams és una biblioteca d'R que es pot descarregar des del repositori oficial CRAN d'R (<https://CRAN.R-project.org/package=exams>) que conté funcions i eines per a generar exàmens de forma automatitzada. Per a confeccionar els qüestionaris es requereix de sintaxis d'R per a fer els càlculs, taules, figures, etc., llenguatge Markdown per a escriure el text formatejat i

algunes instruccions pròpies del paquet exams per a definir el tipus de pregunta (tipus test amb una o més respostes correctes, numèrica, o una barreja de les dues “cloze”). Les preguntes tipus “cloze” permeten, a la pràctica, plantejar diferents qüestions sobre un mateix conjunt de dades generades. Un cop confeccionades les preguntes, es generen enunciats amb dades diferents i per tant preguntes i solucions diferents en cada iteració. Aquest conjunt d’enunciats amb les seves respostes correctes i solucions es guarden en un arxiu que té format XML i que es pot incorporar en un banc de preguntes al Moodle. Finalment, des de la plataforma Moodle es crea un recurs, concretament un qüestionari, on es trien quines preguntes en formaran part, opcionalment definir la puntuació per a cada pregunta (per defecte valen igual) i l’ordre de les preguntes (si és prefixat o s’ordenen de forma aleatòria en cada intent).

Encara que el paquet exams permet exportar qüestionaris a altres formats, com per exemple a PDF amb la finalitat de ser impresos, o en format de graella per a exàmens amb preguntes de tipus test per ser corregides fent servir lectors òptics, etc., per a aquest projecte no s’han considerat. I és que s’han descartat les opcions que només permeten generar exàmens amb preguntes tipus test o que no es puguin corregir de forma automàtica.

4. Conclusions

El paquet exams d’R ha permès crear preguntes individualitzades per a l’assignatura de Disseny d’Experiments del grau d’Estadística de forma fàcil i eficient. El retorn per part dels alumnes ha sigut molt satisfactori. I els resultats docents també han sigut positius. Aquests qüestionaris s’han usat per a la part d’avaluació continuada i per a què els alumnes autoavaluïn els seus coneixements al transcurs de l’assignatura.

Al llarg dels últims cursos, diferents professors encarregats d’impartir part de l’assignatura han confeccionat qüestionaris amb exams sense dificultats, tot i no tenir experiència prèvia amb l’eina. Malgrat la seva simplicitat, s’han pogut crear les preguntes adequades per a cada tema i contingut de cada prova parcial.

L’eficàcia de la introducció d’aquests qüestionaris en la docència i avaluació de l’assignatura s’ha demostrat en comentaris positius i bon “feed-back” per part de l’alumnat, una major participació mesurat amb un increment dels alumnes que trien avaluació continuada i també en una pujada en la nota mitjana respecte als cursos anteriors.

D’acord amb alguns comentaris dels alumnes, per a futurs cursos, fora interessant investigar com integrar la possibilitat que l’estudiant pugui escriure un text que no s’avaluï en el qüestionari, com ara el codi d’R que ha fet servir per a resoldre el problema.

Per tot això, creiem interessant continuar amb la confecció de qüestionaris creats amb l’eina exams així com continuar explorant les possibilitats que ofereix aquest paquet i noves funcions que puguin anar apareixent en versions futures.

Per últim, aquesta estratègia de la utilització del paquet exams es podria aplicar també a altres assignatures i/o d’altres graus de caire més pràctic com de fet s’està usant en l’assignatura de Disseny d’Experiments i Anàlisi de Dades (DEAD) del grau de Biomedicina.

REFERÈNCIES BIBLIOGRÀFIQUES

Grün, B., & Zeileis, A. (2009). Automatic generation of exams in R. *Journal of Statistical Software*, 29(10), 1-14.

Zeileis, A., Umlauf, N., & Leisch, F. (2014). Flexible generation of e-learning exams in R: Moodle quizzes, OLAT assessments, and beyond. *Journal of Statistical Software*, 58(1), 1-36.

Zeileis A (2020). [R/exams: A One-for-All Exams Generator - Online Tests, Live Quizzes, and Written Exams with R.](#) Presented at *Why R? Webinar 001*.
Video: [YouTube](#)