



M2.954 - Estadística avançada aula 1

A0 - Introducció a l'estadística

Inici:	Fi:	Solució:	Solució	Qualificació:	Dedicació:
20/02/19	12/03/19	programada		13/03/19	10 %
		13/03/19			

Descripció i enunciat

Benvinguts a l'activitat introductòria del curs d'estadística avançada. En aquesta activitat, l'estudiant es familiaritza amb l'entorn (l'aula, els professors i la resta d'alumnes) i l'assignatura (pla docent). En aquesta activitat també cal instal·lar el programari necessari i repassar els continguts bàsics de l'estadística per tal de poder construir el coneixement i les competències que es desenvoluparan al llarg del curs a partir d'aquestes bases.

Aquesta activitat no és avaluable, però necessària. L'estudiant ha de participar al fòrum per a presentar-se a l'aula i també ha de realitzar aportacions sobre què sap d'estadística, els àmbits d'aplicació, exemples d'aplicació, ja siguin extrets d'articles o de lectures complementàries o bé de casos concrets més propers al món professional o personal de cadascú, etcètera. L'objectiu d'aquestes aportacions és anar entrant en matèria, començar a parlar en la terminologia estadística i compartir els coneixements previs amb els que cadascú arriba a l'assignatura.

A continuació, s'expliquen els detalls del que cal fer concretament.

1. Familiaritzar-se amb l'entorn

En primer lloc, cal que l'estudiant conegui l'entorn virtual que serà l'eina de comunicació i aprenentatge de l'assignatura. Cal que conegueu les activitats que es proposen, la seva planificació i calendari, el pla docent i el fòrum introductori i fòrums de dubtes.

2. Llegir el Pla Docent

A l'aula virtual, trobareu el pla docent. Cal que el llegiu amb deteniment per conèixer els aspectes essencials de l'assignatura. Un cop llegit, heu de conèixer i comprendre aquests aspectes:

- Què aprendré en aquesta assignatura?
- Quins són els objectius i competències de l'assignatura?
- Quines activitats he de fer?
- Com s'avalua l'assignatura?
- Què succeeix si no lliuro a temps?
- He de lliurar totes les activitats?

Si algun d'aquests aspectes no està clar, podeu revisar el pla docent o formular dubtes al fòrum.

3. Llegir els recursos de l'activitat

La lectura obligatòria associada amb aquesta activitat és el document titulat "Introducció a l'estadística" que està disponible en l'apartat de recursos de l'aula. Hi ha d'altres lectures recomanades que podeu trobar llistades al llarg de l'activitat.

4. Participar en el fòrum

Heu d'accedir al fòrum, com a mínim una vegada, per a realitzar una presentació. En aquesta, podeu indicar aquests temes (és orientatiu):

- Nom i cognoms, professió
- Quines expectatives teniu a l'assignatura
- Què sabeu sobre estadística

Aprofundint una mica més, us recomanem que parleu una mica més en concret sobre l'estadística. Per exemple:

- Algun exemple aplicatiu de l'estadística que heu llegit o proper al món professional
- Alguna eina o metodologia que coneixeu i/o hi heu treballat
- Si l'estadística està present a les vostres vides o no
- Si us agrada l'estadística i el per què
- La importància de l'estadística en el Data Science
- Quines aportacions fa l'estadística en el marc del Data Science

Com veieu, és un fòrum molt obert amb el què es pretén que comenceu a parlar amb terminologia estadística i compartir el que sabeu (o el que no sabeu i voldríeu saber) sobre l'estadística. Al llegir el document d'introducció a l'estadística tindreu idees o preguntes a realitzar en el fòrum.

5. Instal·lar el programari R

El programari que usarem a l'assignatura és el llenguatge R, amb l'entorn RStudio. El document complementari següent us ajudarà a seguir els passos necessaris per a instal·lar el software:

Instalar R.pdf

Aquest document explica com instal·lar R i els aspectes bàsics per familiaritzar-vos amb l'entorn. Seguiu les instruccions del document per a instal·lar el programari. Es recomana que us familiaritzeu amb l'entorn RStudio, seguint les instruccions d'aquest document i realitzant els exercicis que us proporcionem. Així mateix us recomanem que instal·leu i useu el RMarkdown, tal com s'explica més endavant.

Trobareu aquest document a l'apartat "Continguts i Recursos" d'aquesta activitat.

6. Realitzar exercicis introductoris amb R per a conèixer el llenguatge

Per a iniciar el curs amb una base sòlida, cal que tothom conegui i aprengui quins són els tipus de dades en R i sàpiga manipular-los. Per això, us donarem uns exercicis per a practicar i unes lectures recomanades. Són exercicis d'anivellament per als que ja heu treballat anteriorment en R. I pels que no ho conegueu, també us serviran per adquirir els coneixements necessaris. Hi ha també lectures recomanades que us ajudaran a realitzar aquests exercicis.

Concretament, us recomanem que seguiu llegiu el següent document (secció Data Objects):

https://cran.r-project.org/doc/contrib/Lam-IntroductionToR_LHL.pdf

i que consulteu els mòduls:

IntroduccionR.zip

Posteriorment, realitzeu els exercicis:

RExercises.zip

que teniu disponibles a Continguts i Recursos.

7. Realitzar un informe dinàmic usant RMarkdown

Un requisit de l'assignatura és que aprengueu a usar l'RMarkdown per a generar informes dinàmics. Un informe dinàmic és un document text on podeu incloure codi R, amb els resultats de la seva execució. Són molt pràctics per als estudis de Data Science, perquè combina el text escrit i formatejat amb el codi R. És ideal per a generar informes tècnics. El seu ús habitual simplifica molt l'elaboració d'informes. Per a saber com es fan els informes dinàmics, us posem un exemple amb el codi .Rmd i la seva sortida. L'exemple us servirà d'orientació per a generar els vostres propis informes.

L'exemple que us proporcionem és:

RMarkdown.zip

El trobareu a l'apartat de Continguts i Recursos d'aquesta activitat. Si descomprimiu el fitxer, veureu dues carpetes:

- source: hi ha el fitxer font .Rmd i els fitxers necessaris per a interpretar aquest fitxer
- output: hi ha dos fitxers de sortida PDF i HTML, que és la sortida que genera el fitxer .Rmd a l'executar-se.

Us recomanem que després d'instal·lar l'RMarkdown carregueu el fitxer .Rmd que està al source i proveu d'editar-lo i modificar-lo per tal de familiaritzar-vos amb aquesta eina.

Objectius i competències

Els objectius de l'activitat introductòria són:

- Llegir el pla docent
- Familiaritzar-se amb l'entorn (aula, alumnes, professors)
- Conèixer els recursos disponibles de l'assignatura
- Anivellar els coneixements mínims d'estadística necessaris
- Instal·lar el programari necessari
- Llegir el document introductori "Introducció a l'estadística" on s'explica una introducció a l'estadística en el context del Data Science
- Presentar-se en el fòrum
- Realitzar aportacions al fòrum
- Realitzar una activitat voluntària d'introducció a l'entorn i llenguatge R
- Aprendre a fer informes dinàmics usant l'RMarkdown

Continguts i recursos

Els recursos disponibles associats d'aquesta activitat són:

1. Mòdul introductori " **Introducció a l'estadística** ". És el text introductori del curs d'estadística avançada i centra el context de l'estadística en el marc de l'anàlisi de dades. El trobareu a l'apartat Recursos de l'aula.

2. **Moduls R** : 5 mòduls d'introducció a la programació en R. Els trobareu a l'apartat Recursos d'aprenentatge (més avall en aquesta pàgina).
3. **Instalar R.PDF** : document que explica una introducció breu a l'R i com instal·lar el software necessari. (document adjunt en aquest apartat).
4. **RExercises.zip** : conté el fitxer Rmd amb els exercicis de tipus de dades R, el fitxer de sortida html i el fitxers de dades per executar les proves (document adjunt en aquest apartat).
5. **Rmarkdown.zip** : conté un exemple il·lustratiu de com generar informes dinàmics (document adjunt en aquest apartat).

A més, trobareu en aquesta activitat i en les següents mòduls d'un curs d'estadística bàsica que us seran d'utilitat per repassar conceptes i establir les bases a partir de les quals treballarem. Aquests mòduls estan disponibles des de l'apartat Recursos d'aprenentatge de l'aula. A continuació es llisten aquests mòduls juntament amb l'ordre natural de lectura :

1. Estadística descriptiva
2. Mostreig
3. Probabilitat
4. Variables Aleatòries
5. Teorema del límit central
6. Intervals de confiança
7. Contrastos d'hipòtesis
8. Contrast de dues mostres
9. Contrast de variàncies

Recursos de aprendizaje

Mòduls d'introducció al llenguatge i entorn R

1. L'entorn estadístic R

 PDF

2. Àlgebra lineal i càlcul amb R

 PDF

3. Anàlisi de dades i estadística descriptiva amb R i R-Commander

 **PDF**

4. Distribucions de probabilitat i inferència estadística amb R-Commander

 **PDF**

5. Models de regressió i anàlisi multivariant amb R-Commander

 **PDF**

Curs d'estadística bàsica

1. Estadística descriptiva

 **PDF**

2. Mostreig

 **PDF**

3. Probabilitat

 **PDF**

4. Variables aleatòries

 **PDF**

5. Teorema del límit central

 **PDF**

6. Intervals de confiança

 **PDF**

7. Contrast d'hipòtesis

 **PDF**

8. Contrast de dues mostres

 **PDF**

9. Contrast de variàncies

 **PDF**

10. Regressió lineal simple

 **PDF**
