



NÚMERO 41 | ENERO DE 2020

#### En este número...

- Mensaje del Presidente
- Comentario Editorial
- Nueva página web
- Coloquios AME
- Clasificación de documentos usando R
- Anuncios
- Eventos Nacionales
- Eventos Internacionales
- Junta Directiva

## Colaboradores del boletín

Permanentes:

Presidente de la AME:

Luis Enrique Nieto Barajas

lnieto@itam.mx

Editor: Asael Fabian Martínez

fabian@xanum.uam.mx

Co-editora: L. Leticia Ramírez

leticia.ramirez@cimat.mx

### En este número:

- Andrés Christen
- Carlos E. Rodríguez
- Michelle Anzarut
- Teresa Ortiz

# Mensaje del Presidente



A finales del año pasado renovamos la junta directiva de nuestra asociación. Me gustaría comenzar agradeciendo a los miembros de la mesa directiva saliente: Isadora Antoniano, Aurora Montano, Lizbeth Naranjo, Eliud Silva (vocales) y a Gabriel Rodríguez Yam (expresidente) por su labor en la Asociación. Al mismo tiempo quiero darle la bienvenida a los miembros entrantes: Michelle Anzarut, Teresa Ortiz, Carlos E. Rodríguez y Silvia Rodríguez (vocales) y a Andrés Christen (vicepresidente)

por haber aceptado formar parte de la Junta Directiva.

Hemos decidido comenzar el año con una nueva cara. Creamos una nueva página web de la Asociación, la cual se encuentra en la misma dirección que la anterior: amestad.mx. Los invito a que la consulten y espero que sea de su agrado. Agradezco a Teresa Ortiz y a Michelle Anzarut por el trabajo y esfuerzo puestos para la creación de esta nueva página. Más adelante ellas nos relatarán cuales fueron las motivaciones y los objetivos buscados con este cambio.

Como parte de las nuevas iniciativas, hemos creado una modalidad de eventos a los que llamamos "Coloquios AME". Los coloquios consisten en seminarios o talleres impartidos por miembros de la Asociación y apoyados con fondos de la misma. Su objetivo es contribuir a la difusión de la Estadística en México y cualquier miembro activo puede solicitar la realización de un Coloquio en su institución. Más abajo se indicarán los detalles.

Este año 2020 (año par), estará lleno de actividades para la comunidad estadística en México. Tendremos nuestro 34 Foro Nacional en el Centro de Estudios Municipales y Metropolitanos de la FES Acatlán, la Segunda Escuela AME también en las instalaciones de la FES Acatlán, el premio FAO a las mejores tesis en estadística a niveles licenciatura y maestría, y muchos otros eventos en donde la AME colaborará en su organización.

Finalmente quiero agradecer a Carlos E. Rodríguez por su labor como editor del Boletín durante los dos años anteriores y le doy la bienvenida a Fabian Martínez como nuevo editor.

# **Comentario Editorial**

Asael Fabian Martínez y Leticia Ramírez

Esta primera edición del 2020 está dedicada principalmente a dar a conocer las iniciativas de la nueva presidencia. Les invitamos a participar.

De la misma manera, los editores hacemos un llamado a los miembros de la comunidad que deseen realizar contribuciones para el Boletín, ya sea en forma de noticias, eventos o reseñas de eventos, libros o artículos, para que nos las hagan llegar a la dirección de correo fabian@xanum.uam.mx. También son bienvenidas todas las sugerencias que tengan el propósito de enriquecer esta publicación.

Regresar al contenido

# Nueva página web

Michelle Anzarut y Teresa Ortiz

En este mes, la AME estrenó nueva imagen de su página web amestad.mx. La motivación principal fue buscar claridad en la experiencia de navegación en cualquier dispositivo, celular, tablet o computadora. Además, la optimización y mantenimiento se hará de manera constante, permitiendo a los usuarios disfrutar de una página actualizada con las últimas noticias relevantes. De este modo, los socios aprovecharán los beneficios de pertenecer a la AME de forma más sencilla y esperamos que los nuevos visitantes decidan participar en la Asociación.

Visítala para conocer la historia de la AME, los próximos eventos, para leer el Boletín, concursar por el premio "Francisco Aranda Ordaz" o para que la AME te ayude a organizar un Coloquio en tu institución. Además, entre los recursos podrás encontrar la bolsa de trabajo, donde puedes subir tu oferta laboral o leer de otras ofertas.

iEsperamos que te guste!

Regresar al contenido

# **Coloquios AME**

Luis Enrique Nieto

Algo de historia

Durante muchos años, la AME organizó una serie de conferencias llamadas "Conferencias Bimestrales" que tenían como propósito contribuir a la difusión de la Estadística y fortalecer las relaciones entre los diferentes grupos activos de estadísticos en México. La mesa directiva de la AME asistía a estas conferencias y aprovechaban para reunirse y tomar decisiones. En la planeación de estas reuniones, se puso un énfasis especial en descentralizar las actividades de la AME y se procuró que las conferencias se llevarán a cabo en diversas instituciones a lo largo y ancho de país.

En el 2009 se crearon las "Conferencias Itinerantes" cuyo objetivo seguía siendo el de difundir la estadística en México, pero no se tenía la presencia de toda la mesa directiva de la asociación.

Nueva iniciativa

La Asociación Mexicana de Estadística, como parte de sus objetivos, contempla la difusión de la Estadística y el fortalecimiento de los diferentes grupos activos de estadísticos en México. En este marco, la mesa directiva actual propone la creación de los "Coloquios AME" en alguna de dos modalidades: seminario o taller, con las siguientes

### BASES

- 1. Cualquier miembro vigente de la AME puede solicitar un "Coloquio AME" dentro de su institución.
- 2. La AME designará un presupuesto máximo de \$5,000 pesos para cubrir los gastos de traslados (avión o autobús) de un expositor, habiendo un tope de cuatro coloquios máximo por año.
- 3. La institución receptora se compromete a cubrir los gastos de traslados locales, alimentación y hospedaje (en caso de ser requerido) del expositor.
- 4. El solicitante o su institución pueden elegir libremente tanto el tema como el expositor. El tema debe de estar relacionado con la Estadística. Si así lo desean, la AME puede ayudar en la elección del expositor.
- 5. El coloquio se tiene que llevar a cabo en alguna institución del país.
- El seminario o taller tendrá una duración máxima de tres horas.
- 7. El expositor deberá usar los primeros minutos de su presentación para agradecer y presentar a la AME.
- 8. La AME se reserva el derecho de decidir apoyar o no una solicitud.

El formato para solicitar la realización de un Coloquio se puede consultar en la página web (amestad.mx) en la sección de "Participa". Regresar al contenido

## Clasificación de documentos usando R

Carlos E. Rodríguez

El objetivo de esta nota es presentar de manera muy breve las herramientas y funciones necesarias del paquete estadístico R, para clasificar documentos en formato PDF. La clasificación se realizará basándonos en las similitudes del texto entre los distintos documentos.

#### Un caso sencillo

Una buena estrategia para aprender a resolver un problema es empezar trabajando un caso sencillo e ir avanzando paso a paso. Obviamente, el caso más complicado presentará muchos retos adicionales pero para esta nota la idea principal, así como las herramientas básicas son lo importante. Entonces, primero necesitaremos varios documentos en PDF y que a priori, de forma sencilla podamos clasificar, por ejemplo por su título. En nuestro caso, buscamos películas para niños en Wikipedia y salvamos cada una de las páginas encontradas en formato PDF, los archivos que generamos fueron:

```
setwd("/Users/carloserwin/Desktop/Movies/") # Directorio con los PDFs
   files <- list.files(pattern = "pdf$")</pre>
   files # listamos los archivos con formato PDF
    [1] "Cars 1.pdf"
                                           "Cars 2.pdf"
    [3] "Cars 3.pdf"
                                          "Finding Dory.pdf"
    [5] "Finding Nemo.pdf"
                                          "Frozen II.pdf
    [7] "Frozen.pdf"
                                          "How to Train Your Dragon 1.pdf"
    [9] "How to Train Your Dragon 2.pdf" "How to Train Your Dragon 3.pdf"
   [11] "Incredibles 1.pdf"
                                          "Incredibles 2.pdf"
   [13] "Kung Fu Panda 2.pdf"
                                          "Kung Fu Panda 3.pdf"
   [15] "Kung Fu Panda.pdf"
                                          "Madagascar 1.pdf"
11
   [17] "Madagascar 2.pdf"
                                           "Madagascar 3.pdf"
12
   [19] "Monsters 1.pdf"
                                           "Monsters University.pdf"
   [21] "Shrek 1.pdf"
                                           "Shrek 2.pdf'
15
   [23] "Shrek 3.pdf"
                                           "Shrek 4.pdf"
   [25] "Toy Story 1.pdf"
                                           "Toy Story 2.pdf"
   [27] "Toy Story 3.pdf"
                                           "Toy Story 4.pdf"
17
18
19
   length(files)
   [1] 28
```

Es importante mencionar que todos nuestros documentos están en inglés y que además (de suma importancia) los PDF que obtenemos al realizar el proceso descrito anteriormente son documentos en los que podemos seleccionar y extraer el texto. Recordemos que hay documentos en PDF en los que el texto esta embebido por medio de una imagen y es imposible extraerlo simplemente copiando y pegando. En este último caso, lo que suele hacerse, para extraer el texto, es utilizar herramientas de reconocimiento óptico de caracteres (OCR). Esto no se discutirá en esta nota pues va más allá de los objetivos que se persiguen, pero los interesados

pueden buscar la librería tesseract de R <sup>1</sup>.

Para este conjunto de 28 documentos la clasificación es muy obvia, en el primer nivel/grupo deberíamos tener películas de Pixar en un grupo y en otro de DreamWorks, además para el segundo nivel, de cada estudio hemos incluido películas que son secuelas (o pre-cuelas) la una de la otra. Por ejemplo, se incluyeron las 4 películas de Shrek que deberían estar clasificadas en el mismo grupo.

#### Librerías necesarias

Necesitaremos instalar y cargar las siguientes librerías:

- tm: Esta librería brinda un marco para realizar aplicaciones de minería de texto en R. La estructura principal para administrar documentos en la librería tm es el llamado Corpus, que es el texto de los documentos en el que se basará el análisis.
- pdftools: Esta librería nos ayudará a extraer texto de documentos en formato PDE
- wordcloud2: Esta librería permite generar una nube de palabras muy vistosa e interactiva en html. Una nube de palabras es exactamente eso, con la característica de que las palabras que se repiten con mayor frecuencia en el texto, tienen un mayor tamaño. Adicionalmente, se puede controlar el color y formato de las palabras, así como la forma de la nube.

Cargamos las librerías:

```
library(tm)
library(pdftools)
library(wordcloud2)
```

# Extrayendo el texto y generando el Corpus

A partir de los documentos en PDF, necesitaremos extraer el texto y guardarlo en un Corpus, esto se hace de la siguiente manera:

```
corp <- Corpus(URISource(files), readerControl = list(reader = readPDF))
ndocs <- length(corp)
ndocs # documentos leidos
[1] 28</pre>
```

En R, el Corpus es una lista de listas. En el primer nivel, se tienen los documentos, y en cada documento, la primera entrada tiene el texto y la segunda guarda información del

<sup>&</sup>lt;sup>1</sup>Esta librería proporciona enlaces a Tesseract; un potente motor de reconocimiento óptico de caracteres (OCR) que puede reconocer texto en más de 100 idiomas. Esta herramienta es altamente configurable para ajustar los algoritmos de detección y obtener los mejores resultados posibles.

texto de ese documento, por ejemplo: corp[[1]][[1]] contiene el texto del primer documento (Cars 1.pdf), mientras que corp[[1]][[2]] guarda el título del documento, formato, etc. Si se ejecuta corp[[2]][[1]] se obtiene el texto del segundo documento (Cars 2.pdf), etc.

# Homologación y limpieza

Para clasificar los documentos usaremos frecuencias, básicamente construiremos una tabla de contingencia de documentos (renglones) y palabras (columnas). Para calcular frecuencias, conviene "homologar" todo el texto, así que se realizarán los siguientes pasos:

- 1. Eliminar espacios adicionales en el texto.
- 2. Convertir todo a minúsculas.
- 3. Eliminar signos de puntuación.
- 4. Eliminar números.
- 5. Eliminar palabras que se usan con mucha frecuencia, pero que no nos ayudarán a discriminar o diferenciar los documentos. En inglés, se les llama"stop words" y algunos ejemplos son: I, you, me, they, her, their, or, etc.
- 6. Reducimos las palabras a su forma básica (en inglés "steamming"). Esta reducción es el proceso de reducir una palabra a su raíz eliminando sufijos y prefijos. Por ejemplo, las palabras studies, studied, studying, se cambiarán por la palabra study.
- 7. Eliminar palabras que en nuestra opinión no aporten nada al análisis (proceso iterativo). En este punto, se realiza una visualización rápida del texto y eliminamos (o cambiamos) palabras que en nuestra opinión nos ayuden a diferenciar mejor entre los textos de cada documento.

Los primeros 6 puntos, se hacen de la siguiente manera:

```
corp <- tm_map(corp, stripWhitespace)
corp <- tm_map(corp, content_transformer(tolower))
corp <- tm_map(corp, content_transformer(removePunctuation))
corp <- tm_map(corp, content_transformer(removeNumbers))
corp <- tm_map(corp, removeWords, stopwords("english"))
corp <- tm_map(corp, stemDocument)</pre>
```

Estos pasos pueden modificarse u omitirse, pero cuando se hace minería de texto la homologación es un paso fundamental en el análisis.

#### **Análisis**

La tabla de contingencia se obtiene mediante las siguientes instrucciones

Para explorar la tabla de contingencia, convertimos la estructura que arroja DocumentTermMatrix a una matriz. Así, podemos ver, por ejemplo, los primeros 15 renglones y 6 columnas:

```
M <- as.matrix(dtm)
o <- order(sM <- colSums(M), decreasing = TRUE)</pre>
M <- M[,o]
M[1:15, 1:6]
                                   disney
                                           frozen shrek pixar dragon panda
  Cars 1.pdf
                                        27
                                                       0
                                                            49
                                                                     0
                                                 0
                                                                            0
  Cars 2.pdf
                                        26
                                                            43
  Cars 3.pdf
                                        18
                                                 0
                                                       0
                                                            24
                                                                     0
                                                                            0
  Finding Dory.pdf
                                        36
                                                                            0
                                                 3
  Finding Nemo.pdf
                                        23
                                                 2
                                                             38
                                                                            0
                                                       3
                                                                     0
  Frozen II.pdf
                                        69
                                                       0
                                                             5
                                               166
                                                                            0
                                               552
                                                                            0
  Frozen.pdf
                                       276
                                                       0
                                                             10
  How to Train Your Dragon 1.pdf
                                        2
                                                       2
                                                                   148
                                                0
                                                             1
                                                                            1
  How to Train Your Dragon 2.pdf
                                                       0
                                                                   185
                                         1
                                                 1
                                                             0
                                                                            1
  How to Train Your Dragon 3.pdf
                                         0
                                                                   140
                                                                            0
                                                             0
  Incredibles 1.pdf
                                        24
                                                 0
                                                       3
                                                             49
                                                                     0
                                                                            0
  Incredibles 2.pdf
                                        18
                                                            33
                                                                            0
                                                                     0
  Kung Fu Panda 2.pdf
                                         0
                                                 1
                                                       0
                                                                     2
                                                                          112
                                                             1
  Kung Fu Panda 3.pdf
                                         0
                                                 0
                                                       0
                                                             0
                                                                     6
                                                                          114
  Kung Fu Panda.pdf
```

Para generar la nube de palabras simplemente creamos una base de datos y corremos la función wordcloud2 de la siguiente manera:

```
mywords <- data.frame(words = names(sM), freq = as.numeric(sM))[o,]
wordcloud2(data=mywords, size= 0.9, color='random-dark')</pre>
```

La nube de palabras se muestra en la Figura 1.



Figura 1. Nube de palabras

Con la tabla de contingencia M podemos generar nubes de palabras para cada película (e.g. vía la instrucción M[1,], se tienen las frecuencias de cada palabra para la película Cars 1), gráficas de barras y más. En particular, para hacer la clasificación lo que se hace es ver a nuestra tabla de frecuencias como  $M = \{f_{i,i}\}_{m \times k}$ , en donde

- *m*, representa el número de documentos,
- k, representa el número de palabras distintas totales,
- $f_{i,j}$ , número de veces que apareció la palabra j-ésima en el documento i-ésimo.

Sea  $f_i$  el vector de frecuencias para cada una de las palabras distintas del documento i-ésimo, i.e.

$$f_i = (f_{i,1}, f_{i,2}, \dots, f_{i,k}).$$

DATOS NÚMERO 41, ENERO DE 2020

Definimos una matriz de distancias  $D = \{d_{i,h}\}_{mxm}$ , entre los renglones de M, en donde

$$d_{i,h} = ||f_i - f_h|| = \sqrt{\sum_{r=1}^{k} (f_{i,r} - f_{h,r})^2}$$

Con la matriz *D*, podemos utilizar un método de agrupamiento jerárquico, fijando alguno de los criterios de enlace. El dendograma con la clasificación se muestra en la Figura 2.

Se observa que aunque la clasificación es bastante buena, hay problemas con las películas de Frozen y Toy Story, además no se detectó claramente que había películas de dos estudios distintos (Pixar y DreamWorks). Sería interesante determinar los motivos por los que no se hace esta distinción, pero esto va más allá de los objetivos planteados.

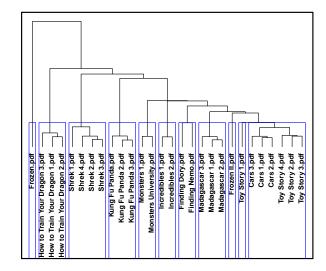


Figura 2. Dendograma

Regresar al contenido

## Anuncios

# Segunda Escuela AME

Carlos E. Rodríguez

Previo al Foro Nacional de Estadística (FNE), los días 7 y 8 de septiembre, se llevará a cabo la Segunda Escuela AME en la FES Acatlán.

Como muchos de ustedes saben, el FNE consistirá principalmente de muchas pláticas cortas (40 minutos aproximadamente) tocando una gran variedad de temas y una sesión de pósters. En cambio, en la Escuela habrá únicamente 4 talleres que durarán 6 horas cada uno.

El público objetivo de las Escuelas son los usuarios de herramientas estadísticas en general, así como estudiantes que buscan incorporarse al mercado laboral.

Los cuatro talleres de la Escuela serán simultáneos, de forma que los asistentes podrán tomar máximo dos. Se buscará que los talleres se enfoquen en temas de interés general para la comunidad estadística, útiles para que el tratamiento de datos sea más sencillo y divertido. Se tiene planeado que uno de los talleres se enfoque en el lenguaje Julia, aunque también se está considerando la posibilidad de incluir un taller de Python. Además de incluir, al menos dos, talleres en donde se utilice el lenguaje R.

En el próximo número del boletín incluiremos el programa completo, así como los costos y otra información relevante.



# 34 Foro Nacional de Estadística: La Estadística en la Multidisciplina

Andrés Christen

En el marco de los 45 años de la Facultad de Estudios Superiores Acatlán, el Foro Nacional de Estadística tendrá lugar en su Centro de Estudios Municipales y Metropolitanos, del 9 al 11 de septiembre. La temática será "la Estadística en la Multidisciplina".

Además de contar con conferencias magistrales, plenaria, sesiones paralelas y una sesión de carteles, en esta ocasión el Foro incluirá dos cursos cortos: "Estadística y Epidemiología" y "Python para Estadística", ambos independientes de la Escuela AME.

Les invitamos a estar pendientes en nuestros medios de comunicación para conocer mayores detalles, como fechas importantes, costos y hospedaje.



# Convocatoria a Contribuciones Libres para el 34 Foro Nacional de Estadística

Se convoca a presentar propuestas de sesiones temáticas y contribuciones libres para el 34 Foro Nacional de Estadística: Estadística en la Multidisciplina, que se llevará a cabo en la UNAM, Facultad de Estudios Superiores Acatlán, los días 9 al 11 de septiembre de 2020. Las sesiones temáticas por contribución las propone un participante e incluyen tres conferencias. Se recibirán propuestas de contribuciones libres

ya sea en formato oral o para la Sesión Magna de Carteles. La fecha límite para recepción de los resúmenes es el 15 de mayo de 2020. En el próximo número indicaremos cómo enviar sus solicitudes. Regresar al contenido

## **Eventos Nacionales**

L. Leticia Ramírez

1. Escuela AME 2020 La Escuela AME surge en 2019 como eventos independientes al Foro anual que consisten de pláticas cortas, sesión de carteles y talleres. Estas actividades buscan formar e informar sobre otras herramientas, modelos y métodos que son de utilidad en un más amplio espectro de la comunidad estadística en el país. El público al que van dirigidas son estudiantes y estadísticos en el sector laboral, público o privado.

Fechas: 7 y 8 de sep. 2020

Lugar: Centro de Estudios Municipales y Metropolita-

nos , FES-Acatlán UNAM, Ciudad de México

Página WEB: A anunciarse

2. Foro nacional de Estadística 2020 Al igual que en sus ediciones anteriores este foro tendrá conferencias plenarias, contribuciones libres, cursos cortos, sesión de carteles y sesiones temática

Fechas: 9 al 11 de sep. 2020

Lugar: Centro de Estudios Municipales y Metropolita-

nos, FES-Acatlán UNAM, Ciudad de México

Página WEB: A anunciarse

Regresar al contenido

# **Eventos Internacionales**

L. Leticia Ramírez

1. SIAM Conference on Uncertainty Quantification (UQ20) La cuantificación de la incertidumbre (UQ) es esencial para proporcionar predicciones computacionales informativas en muchas ramas de la ciencia y la ingeniería. Es un campo interdisciplinario construido sobre una amplia gama de temas básicos matemáticos y estadísticos, y desarrollos algorítmicos y computacionales asociados. UQ20 reunirá a matemáticos, estadísticos, científicos e ingenieros interesados en la teoría, el desarrollo y la implementación de métodos UQ. Si bien se representarán muchos temas, los temas principales de la conferencia incluirán los fundamentos matemáticos y estadísticos de UQ, las conexiones entre los enfoques de UQ basados en datos y basados en datos y las aplicaciones de UQ en las ciencias biológicas y la ciencia climática. El objetivo de la reunión es proporcionar un foro para el fructífero intercambio de ideas entre diversos grupos de expertos técnicos de la academia y la industria, mejorando así la comunicación y contribuyendo a futuros avances en el campo.

Fechas: 24 al 27 Mar. de 2020

Lugar: Garching (cerca de Munich), Alemania

Página WEB: siam-uq20.ma.tum.de

2. International Symposium on Nonparametric Statistics 2020 Nos complace anunciar que el próximo Simposio internacional sobre estadísticas no paramétricas se llevará a cabo en Paphos, Chipre, del 15 al 19 de junio de 2020. En la tradición de las exitosas conferencias ISNPS en Chalkidiki (2012), Cádiz (2014), Aviñón (2016) y Salerno (2018), esta reunión congrega a investigadores nuevos y establecidos para considerar una amplia gama de temas, desde estadísticas no paramétricas y semiparamétricas. El programa incluirá conferencias plenarias, charlas invitadas y contribuidas, y carteles.

Fechas: 15 al 19 Jun. de 2020

Lugar: Paphos, Chipre

Página WEB: cyprusconferences.org/isnps2020

3. Open Problems in Parametric Likelihood-Based Inference El propósito de este evento es reunir a los principales expertos mundiales en inferencia paramétrica clásica basada en la probabilidad para discutir y publicitar importantes problemas abiertos. Las áreas de interés relevantes incluyen los asintóticos de probabilidad (y los asintóticos de orden superior), los fundamentos de la teoría de probabilidad, los problemas en el cálculo y la implementación de los procedimientos de inferencia para probabilidad paramétrica existentes, las áreas de aplicación donde los procedimientos de inferencia

DATOS NÚMERO 41, ENERO DE 2020

basados en probabilidad existentes son inadecuados y otros temas que son principalmente preocupado por la inferencia paramétrica frecuentista basada en la probabilidad. En lugar de conversaciones convencionales que promueven investigaciones recientes, el taller contará con oradores invitados, que representarán a muchos grupos de investigación en diferentes países, discutiendo problemas abiertos desde una perspectiva histórica. También habrá paneles de discusión y sesiones de grupos pequeños. El objetivo del taller es producir un documento con autoridad, intelectual y geográficamente inclusivo, en coautoría de todos los participantes, que explica los problemas abiertos discutidos en el taller y proporciona una evaluación de consenso sobre su importancia. Este documento tiene la intención de servir como guía para nuevos investigadores, editores de revistas y tomadores de decisiones en agencias de financiación, muchos de los cuales pueden no estar familiarizados con el estado de la teoría de probabilidad paramétrica clásica, o qué tipo de resultado constituiría un gran avance. El evento también anunciará la belleza y madurez de este campo fundamental de inferencia estadística, y enfatizará que hay muchos grupos de investigación activos en todo el mundo.

Fechas: 25 al 27 Jun. de 2020

Lugar: St. Louis, EUA Página WEB: Liga

4. ENAR 2020 Spring Meeting La reunión de primavera ENAR 2020 se llevará a cabo en el JW Marriott Nashville del 22 al 25 de marzo. La reunión congrega a investigadores y profesionales de la academia, la industria y el gobierno, conectados a través de un interés común en la biometría.

Fechas: 22 al 25 Mar. de 2020 Lugar: Nashville, Tennessee, EUA

Página WEB: Liga

5. 6th IMA Conference on Mathematics in Defence and Security La ciencia y la tecnología desempeñan un papel cada vez más importante en el apoyo a las industrias de defensa y seguridad. Las matemáticas son fundamentales para estas dos disciplinas, ya que proporcionan un marco para comprender y resolver los problemas variados y complejos que se enfrentan, y para modelar sistemas y escenarios. Estos modelos se utilizan para estimar el rendimiento del sistema, encontrar debilidades en sistemas reales y sugerir mejoras. Este año estamos encantados de que las presentaciones magistrales sean hechas por la profesora asesora científica principal del MOD del Reino Unido, Dame Angela McLean, y el Dr. Ned Allen, científico jefe de Lockheed Martin y miembro principal corporativo.

Fechas: 26 al 26 Mar. de 2020

Lugar: Londres, RU Página WEB: Liga 6. Second Workshop: Emerging Data Science Methods for Complex Biomedical and Cyber Data El taller cuenta con varios oradores eminentes que son pioneros en los campos de la ciencia de datos, incluida la analítica de big data. El taller tiene como objetivo fomentar la investigación colaborativa entre ciencia de datos / estadísticas y otra ciencia disciplinaria con el fin de cumplir con los desafíos científicos más difíciles e importantes basados en datos y modelos, y para mejorar las habilidades muy necesarias de la fuerza laboral de la próxima generación. Hemos sido testigos de la explosión de datos complejos y grandes de diversas disciplinas, redes sociales, tráfico cibernético y el entorno que nos rodea en la última década. Los científicos de datos y los estadísticos han sido bendecidos con tal variedad de datos que nunca antes habían visto, pero también enfrentan muchos desafíos debido a la complejidad y la masividad de dichos datos. El objetivo de este taller se ajusta a la demanda de la sociedad de fomentar la investigación colaborativa entre ciencia de datos / estadísticas y otras disciplinas en ciencia con el propósito de cumplir con los desafíos científicos más difíciles e importantes basados en datos y modelos. Los participantes del Taller aprenderán métodos estadísticos y de ciencia de datos para manejar los datos biomédicos y cibernéticos de enorme complejidad, y ayudarlos a desarrollar el pensamiento analítico, el razonamiento estadístico, las habilidades de comunicación y la creatividad. Representante (s) de IMS en los Comités del Programa: Jie Chen

Fechas: 26 al 27 Mar. de 2020 Lugar: Augusta, Georgia, EUA

Página WEB: Liga

7. International Workshop on Statistical Methods and Artificial Intelligence La inteligencia artificial (IA) se basa principalmente en datos. Utiliza métodos estadísticos a través de las relaciones hombre-máquina durante la generación de datos, la producción de algoritmos y la predicción de resultados. El Taller Internacional sobre Métodos Estadísticos e Inteligencia Artificial será una reunión anual de investigadores en inteligencia artificial, métodos estadísticos, aprendizaje automático y áreas relacionadas. Este taller incluirá (pero no se limitará a) los siguientes temas: 1. Inteligencia artificial 2. Métodos estadísticos 3. Análisis de datos y minería de datos 4. Estadística computacional 5. Aprendizaje supervisado y no supervisado 6. Metodología estadística 7. Bioinformática 8. Estadísticas médicas 9. Aprendizaje profundo 10. Recopilación de datos y aplicaciones 11. Ciencia de datos y tecnología Blockchain 12. Ciencia de datos e inteligencia artificial 13. Ciencia de datos y tecnología Blockchain 14. Estadística matemática 15. Técnicas y aplicaciones de muestreo 16. Software estadístico (R, SAS, Python) El Taller Internacional sobre Métodos Estadísticos e Inteligencia Artificial (IWSMAI) se llevará a cabo del 6 al 9 de abril de 2020, Varsovia, Polonia. El objetivo de IWSMAI es reunir a investigadores, profesores y estudiantes de todo el mundo para presentar sus últimas ideas y resultados de investigación dentro del alcance de IWSMAI 2020. Este taller también proporcionará un entorno ideal para desarrollar nuevas colaboraciones a y conocer expertos en los fundamentos, aplicaciones y productos de los campos mencionados. El taller incluirá presentaciones de ponencias contribuidas, sesiones de pósters y conferencias de vanguardia a cargo de oradores invitados. Los trabajos aceptados IWSMAI 2020 serán publicados por Elsevier Science en la serie en línea de Procedia Computer Science.

Fechas: 6 al 9 Apr. de 2020 Lugar: Warsaw, Polonia

Página WEB: sites.google.com/view/iwsmai

8. Conference on Applied Statistics in Agriculture and Natural Resources La Conferencia sobre Estadística Aplicada en Agricultura y Recursos Naturales reúne a estadísticos de la academia, la industria y el gobierno para discutir ideas y avances en la aplicación de estadísticas para resolver problemas de investigación agrícola. Esta es una conferencia de tres días que consta de un taller, orador principal y una serie de ponencias y presentaciones de pósters.

Fechas: 26 al 28 Apr. de 2020 Lugar: Gainesville, FL, EUA

Página WEB: Liga

9. Frontier Probability Days Representante (s) de IMS en los comités del programa: Davar Khoshnevisan Descripción del programa: Frontier Probability Days 2020 (FPD'20) es un taller regional, que tendrá lugar en la Universidad de Nevada, Las Vegas, Las Vegas, Nevada, del 8 al 10 de mayo. 2020. Su propósito es reunir a matemáticos, tanto regionales como globales, que tengan interés en la probabilidad y sus aplicaciones. FPD tiene como objetivo complementar otras conferencias regionales en Probabilidad que se celebran anualmente en otros lugares de los EE. UU.

Fechas: 8 al 10 May de 2020 Lugar: Las Vegas, NV, EUA

Página WEB: Liga

10. Statistics in the Big Data Era Esta conferencia se centra en el papel y la naturaleza cambiantes de la disciplina de las estadísticas en el momento del diluvio de datos en muchas aplicaciones, y en el éxito creciente de la inteligencia artificial en la realización de muchas tareas de análisis de datos. La conferencia tiene como objetivo reunir a expertos en metodología y teoría estadística para datos complejos y grandes con investigadores centrados en una gama de aplicaciones, desde genómica hasta redes sociales, y brindar oportunidades para que nuevos investigadores aprendan sobre métodos y aplicaciones emergentes. La conferencia también será una ocasión para celebrar el 80 cumpleaños del profesor Peter Bickel. Peter ha pasado su larga y distinguida carrera en el Departamento de Estadística de UC Berkeley, a lo largo de

la cual se mantuvo comprometido con el desarrollo de la teoría y los métodos que arrojan luz sobre las aplicaciones relevantes, un objetivo más relevante que nunca en la era de los grandes datos. Es miembro de IMS y también ha servido a IMS en varias capacidades, incluido el Presidente. Representante (s) de IMS en los Comités del Programa: Liza Levina

Fechas: 27 al 29 May de 2020 Lugar: Berkeley, California, EUA

Página WEB: Liga

11. Mathematical Statistics and Learning La reunión tiene como objetivo congregar a los principales expertos de diversas áreas de estadística matemática y aprendizaje automático que estén interesados en los fundamentos matemáticos de nuestros campos. El tema común de la reunión es el modelado y el análisis estadístico de datos de grandes sistemas complejos, lo que conduce a problemas de alta dimensión y estructurados. Habrá cuatro conferencias especiales matutinas: Francis Bach (INRIA), Liza Levina (Michigan), Luc Devroye (McGill) y Judith Rousseau (Oxford). Representante (s) de IMS en los Comités del Programa: Caroline Uhler, Piotr Zwiernik

Fechas: 2 al 5 Jun. de 2020 Lugar: Barcelona, España Página WEB: www.msl2020.org

12. 6th Stochastic Modeling Techniques and Data Analysis International Conference (SMTDA2020) El objetivo principal de la Conferencia Internacional de Técnicas de Modelado Estocástico y Análisis de Datos (SMTDA) es dar la bienvenida a trabajos, tanto teóricos como prácticos, que presenten nuevas técnicas y metodologías en el área amplia del modelado estocástico y el análisis de datos. Un objetivo es utilizar los métodos propuestos para resolver problemas de la vida real mediante el análisis de los datos relevantes. Además, se promoverá el uso de avances recientes en diferentes campos, como por ejemplo, nuevos métodos de optimización y estadísticos, almacenamiento de datos, minería de datos y sistemas de conocimiento, soporte de decisiones asistido por computación y computación neuronal. Se prestará especial atención a aplicaciones interesantes en ingeniería, producciones y servicios (mantenimiento, fiabilidad, planificación y control, control de calidad, finanzas, seguros, gestión y administración, inventario y logística, marketing, medio ambiente, recursos humanos, biotecnología, medicina, ...)

Fechas: 2 al 5 Jun. de 2020 Lugar: Barcelona, España Página WEB: www.smtda.net

13. **IWAP 2020 (10th International Workshop on Applied Probability)** IWAP es una serie de conferencias bienal con el objetivo de fomentar el intercambio y la fertilización cruzada de ideas sobre probabilidad aplicada. Se realizó en cuatro continentes y atrajo a un buen número de participantes (más de 300) y oradores plenarios

DATOS NÚMERO 41, ENERO DE 2020

distintos en el campo de la probabilidad aplicada y las aplicaciones. La supervisión de las conferencias cuenta con la Junta Internacional de IWAP presidida por el Prof. Joseph Glaz de la Universidad de Connecticut, EE. UU. En junio de 2020, la conferencia IWAP se realizará en Salónica, Grecia, y será organizada por la Universidad Aristóteles de Salónica y presidida por el profesor George Tsaklidis. Para obtener más información, visite la página de inicio de la conferencia: http://iwap2020.web.auth.gr

Fechas: 15 al 18 Jun. de 2020 Lugar: Thessaloniki, Grecia

Página WEB: iwap2020.web.auth.gr

14. Mixtures, Hidden Markov Models and Clustering El objetivo de la conferencia internacional MHC es reunir durante tres días en el Institut de Mathematique d'Orsay algunos investigadores particularmente reconocidos y activos en la comunidad de modelos variables latentes para compartir sus contribuciones más recientes, presentar una comunidad de investigadores jóvenes (Estudiantes de doctorado o postdoctorados) a este tema en constante cambio, pero también y principalmente crean un impulso en el que se discutirán perspectivas interesantes adicionales, como problemas abiertos teóricos, nuevos requisitos de modelado impulsados por las aplicaciones y, en particular, el papel actual de latente modelos variables en la comprensión de estructuras de datos masivas y de alta dimensión.

Fechas: 17 al 19 Jun. de 2020 Lugar: Paris Orsay, Francia

Página WEB: www.math.u-psud.fr/~mhc2020

15. 2020 World Meeting - International Society for Bayesian Este congreso es la continuación de los tradicionales Valencia/ISBA Meetings, celebrados regularmente desde 1979. Estos congresos representan un evento único en donde la comunidad Bayesiana se reúne para presentar los avances recientes y el futuro de la profesión.

Fechas: 29 Jun. al 3 Jul. de 2020 Lugar: Kumming, China

Página WEB: bayesian.org/isba2020-home

16. Bocconi Summer School in Advanced Statistics and Probability: Reproducibility in Data Science La Escuela de verano tendrá lugar del 6 al 17 de julio de 2020, en Villa del Grumello, Lago de Como, Italia. La edición 2020 de la escuela está organizada por la Universidad Bocconi de Milán, en colaboración con la Universidad de Oxford v el Imperial College de Londres, Instructores: Chiara Sabatti (Universidad de Stanford, EE. UU.) Tutoriales Stephen Bates (Universidad de Stanford, EE. UU.) Matteo Sesia (Universidad de Stanford, EE. UU.) La

escuela de verano de dos semanas está abierta a todos los investigadores interesados, pero está especialmente diseñada para estudiantes de doctorado. La fecha límite para las solicitudes es el 7 de abril de 2020.

Fechas: 6 al 17 Jul. de 2020 Lugar: Lake of Como, Italy

Página WEB: bocconi2020.lake-comoschool.org

17. Bernoulli-IMS World Congress in Probability and Statistics 10mo Congreso Mundial de Bernoulli-IMS en Probabilidad y Estadística

> Fechas: 17 al 21 Aug. de 2020 Lugar: Seoul, South Korea Página WEB: wc2020.org

18. 24th International Conference on Computational Statistics (COMPSTAT2020) La 24ta Conferencia Internacional de Estadística Computacional tendrá tutoriales, invitadas y de aportación. La conferencia está patrocinada por la Sección Regional Europea del IASC, y está organizada y organizada por el Departamento de Ciencias Estadísticas de la Universidad de Bolonia.

Las sesiones de tutoría, invitadas, organizadas y aportadas tendrán lugar en paralelo. Una presentación de sesión invitada dura aproximadamente 30 minutos, mientras que una presentación de sesión organizada o contribuida es de 20-25 minutos. Los tutoriales durarán unos 100 minutos.

Fechas: 25 al 28 Aug. de 2020

Lugar: Bologna Italy

Página WEB: www.compstat2020.org

19. 29th International Environmetrics Society (TIES) Conference La conferencia de "The International Environmetrics Society"(TIES) tiene como objetivo fomentar el desarrollo y el uso de métodos y modelos cuantitativos para resolver problemas ambientales. Reúne ideas y perspectivas de investigadores de diferentes disciplinas, como estadísticos, matemáticos, ingenieros, científicos ambientales, informáticos y de datos, para unir el conocimiento y los caminos hacia un mundo más sostenible.

La primera conferencia TIES se celebró en El Cairo (Egipto) en 1989 y desde entonces se ha organizado en diferentes lugares, lo que refleja el carácter internacional de la Sociedad y destaca el enfoque en diversas cuestiones de vanguardia que constituyen las fronteras de la investigación cuantitativa contemporánea en Ciencias Ambientales.

Fechas: 14 al 16 Sep. de 2020

Lugar: Londres, RU. Página WEB: Liga

Regresar al contenido

# Junta Directiva de la AME, 2019-2021

**Presidente:** Luis Enrique Nieto Barajas **Vice-Presidente:** J. Andrés Christen Gracia

**Ex-Presidente:** Ramsés H. Mena **Asistente:** Elida Estrada Barragán

## **Vocales**

Michelle Anzarut Chacalo (Secretaria) Asael Fabian Martínez Martínez Teresa Ortiz Mancera (Webmaster) Paulino Pérez Rodríguez

Lilia Leticia Ramírez Ramírez Luz Judith Rodríguez Esparza (Tesorera) Carlos Erwin Rodríguez Hernández-Vela Silvia Rodríguez Narciso

IIMAS-UNAM, Departamento de Probabilidad y Estadística, Circuito Escolar sin número, Ciudad Universitaria, CP 04510, Delegación Coyoacán, México, CDMX.

correo: amestad@amestad.mx página web: amestad.mx

Twitter: twitter.com/amestadmx

Facebook: clic aquí Regresar al contenido

"Statistics is both a science and an art. It is a science in that its methods are basically systematic and have general application; and an art in that their successful application depends to a considerable degree on the skill and special experience of the statistician, and on his knowledge of the field of application." — Leonard H. C. Tippett