

PARANÁ, 26 SEP 2019

VISTO:

El Expediente N° S01: 3569/2019 UADER\_CYT, referido a la propuesta de la Carrera nueva de Posgrado "Estadística aplicada a la Investigación con R" de la Facultad de de Ciencia y Tecnología cuya Directora es: la Mg. Vaira, Stella Maris; y

CONSIDERANDO:

Que por Resolución CD N° 238/19 del Consejo Directivo de la FCyT- se recomienda la aprobación de la propuesta, la cual cumple con los requisitos esperados de un curso de posgrado y con la normativa vigente (Artículo 4° inc. 4° del anexo II de la Ordenanza 010-06).-

Que la estadística es el nexo común que presentan la mayoría de las investigaciones científicas en las que interviene el tratamiento de datos y la interpretación y predicción de los mismos.

Que dada la importancia que tiene la estadística en la investigación científica, se hace necesario una buena formación para los investigadores, formación que puede ser abordada en las distintas carreras universitarias donde quede plasmado el importante papel que juega esta materia en el desarrollo de las investigaciones y que permita a nuestros futuros investigadores no cometer errores de interpretación originados por el desconocimiento estadístico.

Que este curso tiene por objetivo adquirir un conjunto de conceptos y capacidades y destrezas para aplicar técnicas de análisis estadístico a la resolución de problemas provenientes de diversa índole, en particular de los que se generan desde las ciencias experimentales.

Que a fs. 44/45 la Secretaría de Ciencia y Técnica de la Universidad ha tomado intervención, elevando un informe favorable donde indica que recomienda se dé curso propicio a la presentación bajo la siguiente forma: Denominación del Curso de Posgrado: "Estadística aplicada a la Investigación con R", Director y Docente Dictante: Mg. Stella Maris Vaira, Jefe de Trabajos Prácticos: Lic. Miguel Ángel Fedonczuk, Carga horaria total: 60 (sesenta) horas; Modalidad: Presencial.-

Que la Comisión de Investigación y Posgrado del Consejo Superior, en despacho de fecha 03 de Septiembre de 2019, recomienda la aprobación del curso denominado "Estadística aplicada a la Investigación con R", bajo la dirección de la Mg Vaira, Stella Maris.-

Que el Consejo Superior en su sexta reunión ordinaria llevada a cabo el día 03 de Septiembre de 2019 en el Salón Auditorio "Amanda Mayor" del Rectorado de la Universidad Autónoma de Entre Ríos, resolvió por unanimidad de los presentes aprobar el despacho de Comisión.

Que es competencia de este órgano para resolver actos administrativos en el ámbito de la universidad en uso pleno de la autonomía, según lo normado en el artículo 269° CP E.R. (La Universidad Provincial tiene plena autonomía. El Estado garantiza su autarquía y gratuidad...) y en el artículo 14° incisos a) y n) de la Resolución N° 1181/2001 del Ministerio de Educación de la Nación, Estatuto Académico Provisorio de la Universidad Autónoma de Entre Ríos.-

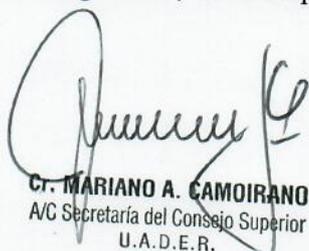
Por ello:

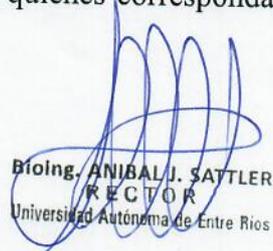
EL CONSEJO SUPERIOR DE LA  
UNIVERSIDAD AUTÓNOMA DE ENTRE RÍOS  
RESUELVE:

ARTÍCULO 1°.- Aprobar el curso de Posgrado denominado "Estadística aplicada a la Investigación con R" de la Facultad de de Ciencia y Tecnología cuya Directora es: la Mg. Vaira, Stella Maris DNI 16.856.806, con una carga horaria total: 60 (sesenta) horas; Modalidad: Presencial, que como anexo único forma parte de la presente.-

ARTÍCULO 2°.- Determinar que la Unidad Académica Responsable: Facultad de Ciencia y Tecnología. UADER, Sede: Oro Verde/Laboratorio de Análisis, Procesamiento y Control de Datos.-

ARTÍCULO 3°.- Regístrese, comuníquese, notifíquese a quienes corresponda y, cumplido archívese.

  
C. MARIANO A. CAMOIRANO  
A/C Secretaría del Consejo Superior  
U.A.D.E.R.

  
Bioing. ANIBAL J. SATTLER  
RECTOR  
Universidad Autónoma de Entre Ríos

ANEXO ÚNICO  
Universidad Autónoma de Entre Ríos  
Facultad de Ciencia y Tecnología  
Curso de Posgrado

1. **Denominación Del Curso De Posgrado:** "Estadística Aplicada a la Investigación con R".

2. **Cuerpo Docente**

Directora y docente dictante del curso: Mg. Stella Maris Vaira (DNI: 16856806).

Jefe de Trabajos Prácticos: Lic. Miguel Ángel Fedonczuk (DNI: 26270570).

3. **Unidad Académica Responsable:** Facultad de Ciencia y Tecnología. UADER.  
Sede: Oro Verde/Laboratorio de Análisis, Procesamiento y Control de Datos

4. **Carga horaria total del curso:** 60 hs.

Carga horaria de actividades teóricas presenciales: 20 hs.

Carga horaria de actividades teórico-prácticas/prácticas presenciales: 20 hs.

Carga horaria de actividades teóricas NO presenciales: hs.

Carga horaria de actividades teórico-prácticas/prácticas NO presenciales: 20 hs

5. **Fundamentación**

La estadística es el nexo común que presentan la mayoría de las investigaciones científicas en las que interviene el tratamiento de datos y la interpretación y predicción de los mismos. El estudio estadístico cobra una especial importancia cuando realizamos investigaciones en diversos campos como la Biología, Ciencias de la Salud, Ciencias Sociales, donde una mala interpretación de los datos, la no adecuada aplicación del método estadístico puede provocar resultados sesgados e incorrectos para la población de interés.

La Estadística es una ciencia derivada de la Matemática que estudia los métodos y procedimientos para: recoger, organizar, y resumir datos; para detectar regularidades,

patrones o tendencias en su comportamiento; para analizar los datos siempre y cuando la variabilidad e incertidumbre sea una causa intrínseca de los mismos, así como de realizar inferencias a partir de ellos, con la finalidad de ayudar a tomar decisiones y, en su caso, formular predicciones.

Uno de sus objetivos fundamentales es utilizar la información suministrada por una parte de la población: la muestra, para hacer inferencias sobre el total de la misma, asociada a una probabilidad de error.

Dada la importancia que tiene la estadística en la investigación científica, se hace necesario una buena formación para los investigadores, formación que puede ser abordada en las distintas carreras universitarias donde quede plasmado el importante papel que juega esta materia en el desarrollo de las investigaciones y que permita a nuestros futuros investigadores no cometer errores de interpretación originados por el desconocimiento estadístico.

El uso de un programa de computación estadístico es importante tanto en la ciencia básica como en la aplicada, por sus posibilidades de automatización de los complejos cálculos estadísticos para el análisis de los datos. El caso específico de R (R Development Core Team) es un programa estadístico y un lenguaje de programación de uso libre, de distribución gratuita y código abierto, desarrollado a partir de un proyecto colaborativo voluntario de investigadores y estadísticos de diversos países y disciplinas. Es un programa basado en comandos, que permite acceder a todos los procedimientos y opciones a través de una sintaxis textual.

R o R Studio es una herramienta de trabajo para el tratamiento y análisis de datos, que ofrece muchas ventajas frente a otras herramientas similares presentes en el mercado:

- R es multiplataforma, libre y abierto.
- Gran número de paquetes disponibles, en muchos casos, programados por los mismos autores de la teoría estadística.
- Una gran comunidad de usuarios y aportantes.
- Ciclo completo de trabajo: implementación de algoritmos, preparación de datos, análisis de resultados y generación de documentación.

## 6. Objetivos

### OBJETIVOS DE APRENDIZAJE

Adquirir un conjunto de conceptos, capacidades y destrezas para aplicar técnicas de análisis estadístico a la resolución de problemas provenientes de diversa índole, en particular de los que se generan desde las Ciencias Experimentales.

Desarrollar conocimientos y habilidades que permitan la participación activa en procesos de investigación. Mostrar las múltiples aplicaciones de las técnicas estadísticas y sus relaciones mediante el análisis de casos y el debate sobre las interpretaciones para cada una de ellas.

## 7. Programa Analítico

**Tema 1.** Introducción. Las variables en la investigación. Clasificación de las variables: cuantitativas y cualitativas. Organización y representación de datos: tablas, gráficos. Medidas de posición y de tendencia central. Medidas de dispersión. Estadística descriptiva bivariada: Tablas de contingencia y diagramas de dispersión. Datos atípicos y faltantes. Aplicaciones. Lecturas de trabajos científicos. Aplicaciones y uso de Software R y RStudio.

Instalación y carga de datos: *install.packages, library, data, class*.

Paquete Análisis Exploratorio de Datos: *paquete EDA, summary, quantile, tapply, plot, barplot, pie, hist, boxplot, table, cov, cor, pairs, scatterplot, scatterplot3d (packages)*.

**Tema 2.** Elementos de muestreo. Población y muestra. Sesgos de muestreo. Muestreo aleatorio simple.

Parámetros y estadísticos. Teorema Central del Límite. Distribuciones de algunos estadísticos. Estimación puntual. Bondad de ajuste. Transformaciones. Aplicaciones.

Kolmogorov-Smirnov en R (con la corrección Lilliefors) para testear la normalidad: *ks.test, lillie.test*

Test de Shapiro Wilks en R: *shapiro.test*

Test gráficos de normalidad: *qqnorm, ppplot, qqplot*.

**Tema 3.** Estimación mediante intervalo de confianza como herramienta para la comparación. Caso de las medias, tanto para muestras dependientes como independientes, caso de las variancias y de las proporciones.

Comandos básicos: *t.test*, *prop.test*, *var.test*, uso de argumentos *conf.int*, *paired*, *conf.level*

**Tema 4.** Pruebas de hipótesis paramétricas para uno y dos grupos. Tipos de hipótesis: Nula y alternativa. Nivel de significancia. Metodología del contraste de hipótesis. Valor P. Contraste para proporciones. Contraste para medias. Contraste para variancias. Introducción al análisis de datos categóricos: pruebas asociadas al estadístico  $\chi^2$ . Medidas de asociación para datos ordinales. Aplicaciones.

Comandos básicos: *t.test*, *prop.test*, *var.test*, uso de argumentos *conf.int*, *paired*, *conf.level*

**Tema 5.** Introducción al análisis de la variancia. Diseño Completamente aleatorizado: efectos fijos y efectos aleatorios. Diseño en bloque. Supuestos. Comparaciones a posteriori. Aplicaciones.

Para trabajar con R: *aov*, *lm*, *TukeyHSD*, *plot(intervals)*, *plot(residuals)*, *shapiro.test(residuals)*, *qqplot(residuals)*, *bartlett.test()*.

**Tema 6.** Regresión lineal simple. Estimación por el método de mínimos cuadrados ordinarios. Inferencia de los coeficientes de regresión. Medidas repetidas: calidad del ajuste. Descomposición de suma de cuadrados para Regresión. Análisis de residuos. Coeficiente de correlación y coeficiente de determinación. Aplicaciones.

Para trabajar con R: *lm()*, *summary()*, *plot()*, *cor.test()*, *hist()*, *rstandard()*,

## 8. Metodología de la enseñanza

El desarrollo de las clases será teórico-práctico, se iniciarán con conceptos y métodos estadísticos, luego se desarrollará un problema completo con software, en el que se incluye la carga de datos, la selección de la técnica y la interpretación de los resultados.

Para las exposiciones se recurre al apoyo de proyecciones (CAÑON), lo que agiliza notablemente el desarrollo de los encuentros. Debido a la temática y al enfoque del curso (eminentemente aplicado), la mayoría de los temas del Programa se presentan inicialmente con un caso práctico a resolver, y se van introduciendo los aspectos conceptuales y métodos estadísticos a medida que se hacen necesarias, con la fundamentación correspondiente.

Para la realización del curso se requiere, como criterio de inclusión, que los alumnos acrediten conocimientos básicos de Estadística e Informática y deberán asistir a clase con su notebook.

### Software R - R Studio.

#### 9. Destinatarios

Graduados universitarios, docentes y profesionales que quieran profundizar en estadística aplicada utilizando un software libre, como lo es R, para analizar, procesar y hacer lectura científica crítica de trabajos publicados o a publicar.

#### 10. Cupos

Cupo mínimo: 20 cursantes.

Cupo máximo: 30 cursantes.

#### 11. Requisitos de aprobación

La evaluación final será mediante un trabajo final individual o grupal, tipo time-home. El trabajo final se entregará vía mail tres a cuatro semanas después de concluido el curso. La devolución del mismo será vía mail y de ser necesario se pautará un encuentro para aclarar dudas. Se exigirá al menos 75% de asistencia.

En el examen se intenta reproducir las situaciones reales de una actividad laboral en Estadística, se deberá elaborar "informes" y dar respuesta a los temas y problemas planteados, es factible consultar cualquier tipo de bibliografía y antecedentes (es decir que los Exámenes son "a libro abierto"), pero se los debe finalizar en un tiempo delimitado previamente definido (el tiempo dedicado al Examen forma parte de la Evaluación).

#### 12. Bibliografía

Mendenhal, W.; Beaver, R.; Beaver, B. (2010). *Introducción a la probabilidad y estadística*. Cengage Learning Editores. Mexico.

Elorza Perez-Tejada, H. (2008). *Estadística para las Ciencias Sociales, del Comportamiento y de la Salud*. Cengage Learning Editores. 3° Edición. Mexico.

Denzin, N.; Lincoln, Y. (2011). *El campo de la investigación cualitativa*. Gedisa Editorial. Barcelona. España.

- Peña, Daniel. (2002). Regresión y diseño de experimentos. Alianza Editorial, S. A. Madrid.
- Agresti, A. (2002). Categorical data analysis. John Wiley & Sons, Inc., Publication. 2nd edition. USA.
- Dalgaard P (2002). Introductory Statistics with R. Springer-Verlang, New York.
- Friendly M (1994). Mosaic Displays for Multi-Way Contingency Tables. Journal of the American Statistical Association, 89, 190-200.
- Kleiber y Zeileis (2008). Applied Econometrics with R. Springer. New York.
- Venables, William N., David M. Smith, and R Development Core Team. An introduction to R., 2002.
- Tan, P. N., Steinbach, M., & Kumar, V. Introduction to Data Mining, 2005.
- Mirkin, B. (2001). Eleven ways to look at the Chi-Squared coefficient for contingency tables. The American Statistician, 55, 111-120.
- Montgomery, D.C. and Peck, E.A. (1992). Introduction to Linear Regression Analysis, Second edition. John Wiley and Sons, Inc.
- Myers, R.H. (1990). Classical and Modern Regression with Applications, 2nd ed. PWS-Kent Publishers, Boston. <http://cran.r-project.org>

