

Proyecto docente Justificación y programa de la asignatura

Objetivos

Los objetivos generales que tenemos en la asignatura son los de conseguir que los estudiantes sean capaces de

- realizar análisis y comprender otros ya realizados,
- interpretar correctamente la dimensión numérica y estadística que nos rodea, a través principalmente de los textos, los medios de comunicación y los discursos,
- utilizar los conocimientos adquiridos en la asignatura para manejarse con soltura en las tres dimensiones de interés: la cotidiana, la profesional y la científica.

A su vez, el primer punto puede ser extendido en cinco momentos, organizados por orden temporal:

1. Identificar correctamente la situación.
2. Tomar una buena decisión sobre qué recurso estadístico aplicar.
3. Aplicarlo (realizar cálculos y/o expresar a un programa de ordenador que lleve a cabo la tarea).
4. Interpretar correctamente el resultado.
5. Comunicar efectivamente.

Tres llamadas de atención

Es importante recordar la función universitaria. Diferentes documentos (desde la Declaración de Educación Superior de la UNESCO de 1998, hasta la LOU) impelen a los miembros de la universidad, especialmente al profesorado, a insertar los conocimientos que se van adquiriendo dentro de la actualidad local y mundial, de tal forma que los estudiantes terminen siendo profesionales y ciudadanos no compartimentados, capaces de comportarse en su entorno con responsabilidad e implicación sociales, participando en las soluciones más que en los problemas. La forma en que las personas aprendemos aconseja también ir practicando los objetivos de aprendizaje insertos en asuntos de interés general que aumentan, además, la motivación. Los acontecimientos de actualidad (como ocurre en estos momentos con la llamada crisis económica, por ejemplo) son un buen motivo para abordar los diferentes elementos de análisis de datos, tanto para generar una mirada estadístico-metodológica, como para interpretar correctamente los números que nos rodean y ser capaces de utilizar las habilidades que estamos adquiriendo para transitar con conocimiento y autonomía por la cotidianidad.

Es importante también remarcar la importancia del sentido común. Resulta harto habitual que las personas se desposean de él cuando se sitúan ante un ordenador y realizan análisis de datos. El diseño de investigaciones y el análisis de los datos no son campos de excepción a la norma general de contar con autonomía de pensamiento, sentido común y capacidad crítica en

continua activación, evitando automatismos y comportamientos irreflexivos. El sentido común y una actitud del tipo “voy a comprender y dominar esto”, son excelentes puntos de apoyo.

Por último, nuestras clases de análisis de datos se imparten en psicología. Ello condiciona totalmente el contenido y, especialmente, la forma de impartirlo. La estadística comparte muchos esquemas con la psicología y ello ayuda a observar la estrecha relación entre ambos campos del saber. Por ejemplo, ambos están muy preocupados y ocupados en los conceptos de normalidad, norma, referencia, etc. Buena parte de la estadística puede observarse como la preocupación por identificar lo común y distinguir lo particular, algo que se encuentra también en la esencia de la ciencia y la profesión de la psicología. Si aprovechamos estas conexiones estrechas, el análisis de los datos se observa familiar, natural para y desde la psicología. Eso aumenta la esperanza de aprender.

Esquema básico y temporalmente ordenado de subobjetivos de análisis

1. Identificar claramente el reto que hay que cubrir: conocer qué valor representa mejor a un conjunto de datos, concluir si existe relación entre dos variables, etc. La identificación se realiza sin aislar el reto de su contexto.
2. Identificar las variables implicadas en la situación, catalogando correctamente su tipo según la clasificación habitual: nominal, ordinal, cuantitativa.
3. Conocer cada variable implicada por separado, utilizando los recursos apropiados según tres criterios: qué objetivos se persiguen, de qué tipo es la variable y cuál es su comportamiento particular.
4. Distinguir cuál es el objetivo final que motiva el proceso de análisis. No todos los objetivos caben en esta asignatura. Los que sí, pueden agruparse en cinco categorías:
 - a) Interpretar un comportamiento grupal:
 - Tabular o representar gráficamente todas las puntuaciones, observando el comportamiento colectivo. Hay que ser capaces de explicar a otras personas cómo se comporta esa variable.
 - Observar si algunas puntuaciones son raras o peculiares y poner en marcha dos procesos: uno para el grueso más homogéneo y otro para las puntuaciones peculiares.
 - Si procede, investigar por qué aparecen esas puntuaciones peculiares, caso de existir.
 - b) Encontrar un valor representativo o característico para el conjunto de datos de una variable:
 - La variable debe ser suficientemente conocida y depurada antes de abordar específicamente este punto.
 - Observar si procede realizar subgrupos y concluir con un valor representativo para cada uno de ellos.
 - Escoger las medidas que van a ser utilizadas para la representación (habitualmente un valor numérico y una medida de cuán bueno es).
 - Observar si procede el valor característico o si lo característico es la variabilidad, de tal forma que cualquier elección de un valor concreto sería equivocada.
 - Distinguir si procede inferencia (estimación por intervalo) o no. Si procede, tomar buenas decisiones sobre la precisión y la seguridad, considerando el objetivo, el contexto y las características de la variable.

c) Interpretar una puntuación individual:

- El valor individual se interpreta en función de su grupo, por lo que el comportamiento grupal debe ser convenientemente conocido y los datos encontrarse depurados (es decir, se debe atravesar antes la fase 4a).
- En función de las características de la variable y del objetivo, escoger el recurso estadístico conveniente para traducir el dato original a una puntuación derivada que facilite la interpretación. Es posible que se requiera más de una puntuación. Tal vez sea recomendable acompañar el valor numérico con una representación gráfica específica.
- El objetivo puede ser construir una tabla o baremo que permita interpretar en el futuro distintos resultados individuales. En este caso, se toman dos decisiones: qué puntuación derivada utilizar y cómo disponer la tabla.

d) Concluir sobre la posible relación entre dos variables:

- Aplicar la fase 4a para cada una de las dos variables por separado.
- Realizar una tabulación o representación gráfica bivariada, de tal forma que se observe la posible relación considerando todos los datos. A la vista de este resultado hay que ser capaces de explicar a otra persona qué está ocurriendo.
- Depurar, si procede. Es posible que se observe:
 - La existencia de algunos comportamientos individuales anómalos que desdibujarían las conclusiones de relación para el grueso.
 - La conveniencia de distinguir entre dos o más grupos de datos, realizando el análisis para cada uno por separado.
 - La existencia de una escala de medida inconveniente, que aconseja algún tipo de transformación para una o ambas variables.
- Cuantificar la relación, es decir, escoger y aplicar un recurso estadístico que represente el grado de relación mediante un número interpretable. Ello implica:
 - Definir el objetivo de relación en términos estadísticos. Por ejemplo: comparación entre dos medias aritméticas.
 - Escoger el recurso adecuado según el objetivo y las características de las variables implicadas. Por ejemplo: aplicar una *t* de Student.
 - Obtener la cuantía e interpretarla. Para la interpretación es posible que deba traducirse el valor original a una escala manejable. Por ejemplo: obtener un valor *V* de Cramer (acotado en 0-1) a partir de una Chi cuadrado.
- Tomar la decisión de si se continúa con el proceso hacia la inferencia o ya es factible concluir sin realizarla. Esto segundo puede ocurrir cuando:
 - No se está trabajando con una muestra sino directamente con la población. En tal caso, el único recurso para concluir sobre la existencia o no de relación es la cuantía del efecto obtenido.
 - La muestra no es aleatoria y por tanto no procede poner en marcha inferencia estadística. Se concluye como en el guión anterior.
 - La cuantía obtenida expresa un efecto insuficiente para concluir que existe relación (por ejemplo, un valor *V* de Cramer muy cercano a cero).
- Si procede, aplicar una PSHN (prueba de significación de la hipótesis nula):
 - Traducir el objetivo de relación a un hipótesis nula.
 - Situar el resultado muestral en un contexto de interpretación (un intervalo de estimación, una distancia estandarizada o un grado de significación).
 - Decidir: la hipótesis nula se rechaza o se mantiene.

- Concluir: traducir la decisión a los mismos términos que el enunciado de investigación.
- e) Realizar un muestreo. Aunque un buen muestreo es algo más complejo de lo que se aprende en la asignatura, es necesario abordar algunos puntos mínimos. Nos situamos únicamente en el contexto de una estimación por intervalo a partir de un muestreo aleatorio simple, y solo considerando medias o proporciones.
 - Escoger un objetivo de análisis (estimación de media o de proporción) o un paquete lo más pequeño posible de objetivos que van a guiar el muestreo.
 - Decidir qué valores se desean para realizar la estimación en términos de:
 - Seguridad. Tomar esta decisión a partir de las consecuencias que se derivan de errar en la estimación y de lo que es habitual en contextos similares.
 - Precisión. Tomar esta decisión teniendo en cuenta el objetivo y contexto, además del valor esperado para la característica implicada (media aritmética o proporción), imaginando un intervalo concreto admisible.
 - Considerar alguna medida de variabilidad para la característica en la población:
 - Tener en cuenta estudios previos similares.
 - Realizar un estudio piloto.
 - Deducir aritméticamente un valor (como, por ejemplo, un valor máximo para la varianza).
 - Calcular el tamaño de la muestra a partir de la información anterior.
- 5. Redactar un texto que describa el objetivo, el proceso, los resultados y las conclusiones, a partir del conocimiento generado, sabiendo utilizar los mejores recursos estadísticos para comunicar, insertar en el texto los resultados generados por ordenador y teniendo en cuenta quiénes son los receptores del texto (la misma persona de cara a su archivo profesional, medios de comunicación, comunidad científica, otros profesionales...). Es posible que el texto sea específico para un objetivo muy concreto, como ocurre en la construcción de un cuadro de muestreo o en un informe breve de conclusiones.

Organización del contenido por unidades

Los temas no son abordados en el orden exacto con que se exponen en el programa, sino que se organizan en sentido pedagógico en las siguientes unidades de aprendizaje, cada una de ellas provista de materiales específicos de teoría y práctica. Antes de estas unidades tiene lugar una sesión de presentación de la asignatura y del sentido del análisis de los datos en la ciencia y práctica de la psicología, así como en la vida cotidiana.

1. Conocer una variable. Efectos de la escala de medida (nominal, ordinal, cuantitativa) en las posibilidades de análisis de datos. Medida de la norma y del error y su relación con intereses psicológicos. Tabulación. Representación gráfica: sentido y concreciones (ciclograma, diagrama de barras, histograma, diagrama de caja y patillas univariable). Representación numérica: sentido, norma o tendencia central (moda, mediana, media) y bondad de la norma o dispersión (% de la moda, MAD, varianza, desviación tipo, coeficiente de variación, variación acotada). Comparación entre normas.
2. Interpretar un caso. Necesidad en la ciencia, en la profesión de la psicología y en la vida cotidiana. Relación entre caso y grupo de referencia. Medidas de posición: partes, puntos de corte e interpretación. Dificultades de interpretación de las medidas de posición.

Distancia estandarizada: sentido; relación entre caso, norma y bondad de la norma; denominaciones y utilización. Propiedades de las distancias estandarizadas. Cambio de escala y utilización en tests psicológicos. Interpretar la distancia entre lo esperado y lo encontrado. Chi cuadrado de Pearson univariable. V de Cramer.

3. Relacionar gráficamente dos variables. Sentido de una relación. Efectos de la variabilidad unitaria. Criterios para la clasificación de técnicas de representación gráfica. Situaciones estándar en la tabla de contingencia, el diagrama de puntos, el diagrama de medias y el diagrama de caja y patillas combinado. Excepciones y situaciones especiales.
4. Cuantificar la relación entre dos variables. Sentido y limitaciones de la cuantificación. Criterios para la clasificación de técnicas para cuantificar relaciones bivariantes. Paramétrico *versus* no paramétrico (primera aproximación). Chi cuadrado de Pearson, coeficiente de correlación lineal simple de Pearson, t de Student-Fisher, t de Wilcoxon, U de Mann-Whitney, análisis de la varianza.
5. Diseñar una muestra. Sentido y limitaciones del muestreo. Vocabulario y símbolos. Modelos de muestreo aleatorios (simple, sistemático, de conglomerados, estratificado, con submuestreo) y no aleatorios (de voluntarios, circunstancial, de cuotas, de bola de nieve). Distribución muestral. Valor esperado y error tipo. Estimadores insesgados.
6. Manejar la curva normal. Sentido e historia de la curva normal. Trascendencia y limitaciones en la utilización de la curva normal. Características y comportamiento. Relación con las distancias estandarizadas. Relación con las distribuciones muestrales. Utilización de tablas.
7. Realizar una estimación estadística. Sentido. Qué se estima. Estimación puntual. Estimación por intervalo (estimador, error de precisión, error de muestreo, riesgo y confianza). Criterios para tomar decisiones y realizar cálculos sobre los elementos de una estimación por intervalo.
8. Decidir un tamaño de muestra. Relación entre los objetivos de la estimación por intervalo y el tamaño de la muestra. Lógica de las expresiones de cálculo. Relación entre el tamaño y otros elementos de una inferencia (tamaño de la población, variabilidad poblacional, seguridad y precisión). Tamaño de muestra para múltiples objetivos. Útiles informáticos.
9. Realizar una prueba de significación de la hipótesis nula. Hipótesis de investigación, estadística y nula. Metáfora del sistema de justicia. Lógica de la compatibilidad. Fases de una PSHN. Alternativas de concreción: estimación por intervalo, distancias directas, distancias estandarizadas y probabilidades. Pruebas de una y dos colas.
10. Inferir en los casos de relación entre variables. Fases en el estudio de una relación entre variables. Conexión entre tamaño de efecto en la muestra e inferencia estadística. Distribuciones muestrales asociadas a los esquemas de relación especificados en la unidad de aprendizaje 4 (cuantificar la relación entre dos variables). Paramétrico *versus* no paramétrico (segunda aproximación). Estimación y prueba de significación de la hipótesis nula. La trampa de la comparación entre probabilidades. Buenos hábitos para inferir relaciones.