

Análisis de datos en psicología I

Apuntes sobre inferencia

Vicente Manzano · 2005

Algunos conceptos necesarios para entender bien estos apuntes se han impartido ya antes en la asignatura de Fundamentos Metodológicos en Psicología. Se refieren a muestra, población, modelos de muestreo y escalas de medida, además de nociones sobre diseño, covariación y causalidad. Así que en estas páginas no vamos a detenernos en ello, aunque sí se harán referencias y, ocasionalmente, alguna explicación.

Se ha procurado que las explicaciones sean muy claras y que estén asentadas en ejemplos concretos con SPSS, incluyendo las salidas del ordenador comentadas. No obstante, es posible que algo quede de difícil comprensión. El esfuerzo empleado para realizar estos apuntes se justifica por la motivación de que sean útiles. Así que se agradecerá cualquier opinión, sugerencia o comentario que vaya orientado a solucionar esos puntos de difícil comprensión, errores o lagunas.

¿Qué es inferencia?

Inferir es "sacar una consecuencia o deducir algo de otra cosa". Así lo define el Diccionario de la Real Academia Española. Y no está mal, se adecua al uso que hacemos en análisis de datos.

En nuestro caso, queremos saber cosas sobre una población, pero ésta es muy grande, inaccesible o de estudio costoso. En definitiva, hemos decidido extraer una muestra, estudiar ésta y, después, *inferir* qué habríamos encontrado de haber trabajado directamente con la población. Por ejemplo, nos interesa conocer qué piensan los habitantes de Sevilla capital con respecto a la última película de Carlos Saura. Para ello, elaboramos un cuestionario de veinte preguntas. La capital de Sevilla cuenta con unos 700 mil habitantes. Lo tenemos difícil y desconfiamos de contar con el control suficiente como para garantizar que aún invirtiendo una gran suma de dinero, consiguiéramos creer en los resultados obtenidos. Así que hemos decidido trabajar con una muestra. Supongamos que está compuesta de 300 personas. Preguntamos a éstas y contamos con sus respuestas en un archivo de datos que hemos analizado y que nos suministra información sobre la opinión *exclusivamente de estas 300 personas*. ¿Qué hacer con las 700 mil de la población, que son las que realmente interesan al estudio?

La solución está en partir de cómo se ha obtenido la muestra y de inferir, a partir de ello y de los resultados obtenidos, qué habríamos encontrado de haber trabajado, en las mismas condiciones, con la población. Es importante resaltar en esto que la inferencia depende del muestreo. Es decir: no se puede utilizar cualquier procedimiento de inferencia, sino aquél o aquéllos que tengan sentido teniendo en cuenta cómo se ha obtenido la muestra. Por ejemplo, si lanzamos un dado al aire, sabemos que pueden salir cualquiera de sus seis caras, con igual probabilidad. Si tenemos que apostar por un resultado, no tenemos información sobre qué cosa es la que ha salido o puede salir en concreto. Pero si sabemos que Gerardo es un tramposo y que tiene una predilección obsesiva por el número 5, y le pedimos que lance un dado al aire, es razonable esperar que el resultado sea, precisamente, el número 5 (nos podemos equivocar, claro, pero ¿A qué apostarías?) Los procesos de inferencia son una especie de "camino de vuelta" con respecto al muestreo, que hace de "camino de ida".

La inferencia que vamos a ver es la que se denomina "estadística". Es una inferencia basada en probabilidades. Para aplicar inferencia estadística es necesario que la muestra sea aleatoria. Si la muestra con la que estamos trabajando es de voluntarios, es la de un grupo de clase en psicología o surge de las personas que me han parecido más propensas a responder de entre las que me he encontrado por la calle un sábado por la mañana, entonces no puedo aplicar inferencia estadística.

¿Cómo se hace inferencia estadística?

Lo primero que se necesita ya se ha dicho: es imprescindible que la muestra sea aleatoria. En esta asignatura (y, en general, en toda la licenciatura) vamos a partir de que se ha recurrido a una muestra aleatoria simple.

Lo siguiente es que tenga sentido hacer la inferencia. Recordemos que se trata de inferir qué habríamos obtenido de haber trabajado con la población en lugar de con la muestra, en las mismas condiciones. En muchas ocasiones, los resultados de la muestra son suficientemente claros como para saber qué pasa en la población y, por tanto, no es necesario hacer inferencia estadística. Imaginemos, por ejemplo, que hemos medido la relación entre el canal de televisión que se ve habitualmente y el número de veces que la persona come fuera de su casa en un año. Al analizar estos datos en la muestra, observamos que todas las personas comen fuera en 26 ocasiones. Y esa cantidad es exactamente la misma en todos los grupos de personas (en el grupo que ve La 2, en el grupo que ve A3, etc.) Ya en la muestra vemos que no hay ningún indicio o sospecha de que pueda haber diferencias, así que carece de sentido hacer ninguna inferencia a la población. Casos tan exagerados como éste no se dan en la práctica. Pero nos sirve como ejemplo para comprender que no siempre hay que hacer una inferencia estadística.

Pues bien, si la muestra es aleatoria y tiene sentido hacer inferencia estadística, entonces se pone en marcha un proceso que parece complicado (vale, lo es) y del que vamos a tener la primera impresión ahora mismo, en el cuadro 1. No te preocupes si lo ves muy abstracto, es la primera visión de bosque y después nos introduciremos en sus árboles.

1. La muestra que hemos obtenido y con la que hemos trabajado sólo es una de las muchas muestras que pueden obtenerse de la misma población.
2. Lo que hacemos es pensar en la población y plantearnos qué resultados se pueden obtener cuando se extraen muestras. Esos resultados se organizan en lo que denominamos una *distribución muestral*.
3. En esa distribución muestral localizamos nuestra muestra. Más bien, lo que hacemos (porque ya veremos que no podemos hacer otra cosa) es calcular una distancia posible entre el resultado en la muestra y el que habríamos obtenido en la población.
4. Con esa distancia, calculamos un intervalo alrededor del valor obtenido en la muestra y afirmamos: "esperamos que el valor de la población se encuentre en este intervalo".

Cuadro 1: esquema de una estimación por intervalo.

En inferencia hay muchos procedimientos y nos podemos plantear muchos objetivos. Lo que vamos a abordar en este documento es un par de esas posibilidades, las más extendidas, utilizadas e importantes en la práctica: realización de estimaciones mediante la construcción de *intervalos de confianza* y toma de decisiones mediante la *prueba de significación de la hipótesis nula*. Lo que se ha descrito en el esquema anterior es el guión de una estimación por intervalo de confianza, pero puede servir de base para tomar decisiones con la prueba de significación de la hipótesis nula. Así que vamos a seguir en ese mismo orden: primero aprenderemos a estimar mediante intervalos y, después, a decidir mediante pruebas.

Antes de seguir, conviene recordar dos términos. Sabemos que el análisis de datos nos permite resumir la información en tablas, dibujar gráficas y calcular índices que surgen de la aplicación de fórmulas o algoritmos. Pues bien, cuando los índices se calculan con los datos de la muestra se denominan *estadísticos*, mientras que al calcularse con los datos de la población reciben el nombre de *parámetros*. El término parámetro se utiliza también para otros conceptos, pero tal como se ha expuesto en este párrafo es como se encuentra más extendido su uso y vamos a seguirlo aquí.

Entonces, la inferencia consiste en deducir el valor de un parámetro a partir de un estadístico.

Un último aporte de vocabulario y simbología. Para no confundirnos entre estadísticos y parámetros, se suele utilizar una fuente latina para los primeros y griega para los segundos. En el cuadro 2 se encuentran los símbolos más comúnmente utilizados.

	índices		
	media	proporción	des. tipo
muestra	\bar{X}	p	S
población	μ	π	σ

Cuadro 2: símbolos para muestras y poblaciones.

Y ánimo, que esto no es fácil, pero sí posible.

La distribución muestral

La distribución muestral es la distribución de todos los valores que puede suministrar un estadístico cuando se calcula en todas las muestras posibles desde una población. No se trata exactamente de todas las muestras posibles, porque se refiere sólo a las que tienen un tamaño concreto (por ejemplo, $n=45$) y se han obtenido con el mismo procedimiento (por ejemplo, el aleatorio simple). Veamos un ejemplo.

Tenemos una muestra aleatoria simple de 40 personas a las que hemos preguntado si fuman o no. Hemos calculado el porcentaje en la muestra (estadístico) y el resultado ha sido que fuman 5, es decir, el 12,5%. Esto podríamos hacerlo con todas las muestras aleatorias simples de 40 personas de esa misma población. En una ocasión, obtenemos 13%, en otra, 12,8%, etc... Tendremos entonces tantos resultados como muestras posibles. Todos esos resultados (12,5; 13; 12,8; ...) son datos, por lo que podemos hacer con ellos lo mismo que con cualquier conjunto de datos:

1. Organizarlos en una tabla o gráfico, generando una distribución, que se denomina muestral. En el ejemplo, se trata de una distribución muestral de proporciones (podría ser de medias, de varianzas, de diferencias entre medias, etc.)
2. Calcular su media aritmética, que aún siendo una media como cualquier otra, recibe un nombre especial por referirse a una distribución muestral: esperanza matemática o valor esperado. Así, la media de la distribución muestral de proporciones se llama *valor esperado de la proporción*.
3. Calcular su desviación tipo, que del mismo modo, recibe una denominación especial: *error tipo*. La desviación tipo de la distribución muestral de proporciones recibe el nombre de error tipo de la proporción.

En la práctica no se construyen distribuciones muestrales reales, sino que se deducen. Esto es así porque requieren un gran esfuerzo. Con poblaciones muy pequeñas ya se obtiene un número prácticamente infinito de muestras posibles. Así que algunas personas se han dedicado a realizar deducciones sobre distribuciones muestrales teóricas. Y esas deducciones permiten:

- Conocer la forma de la distribución, de tal manera que se sabe, por ejemplo, cuán probable es encontrar determinados resultados muestrales.
- Saber en qué condiciones un muestreo real genera una distribución muestral teórica.

En definitiva, pues, no vamos a calcular en la práctica una distribución muestral, sino que vamos a suponer que es de determinado tipo si se cumplen unas condiciones concretas. En función del muestreo realizado y de las características de la variables que estamos estudiando, podremos suponer que la distribución muestral es de un tipo determinado y conocido. Gracias a ello, podremos realizar el paso 3 del esquema de inferencia que has visto en el cuadro 1. Vamos a comenzar abordando el proceso con la distribución muestral más famosa, extendida y conocida: la normal.

La distribución normal

Algunos teóricos, hace ya siglos, la generaron y estudiaron. Algunos de ellos fueron Laplace y Gauss. Por otro lado, al representarla gráficamente se obtiene una curva que tiene forma de campana. Por ello se le conoce con muchos nombres: curva normal, ley normal, distribución normal, campana de Gauss, curva de Laplace, función de Laplace-Gauss, etc. La figura 1 muestra un ejemplo concreto.

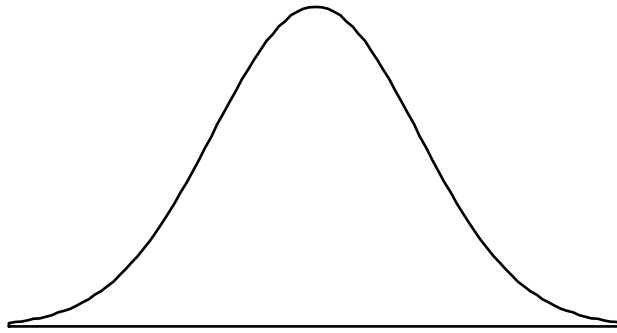


Figura 1: curva normal.

Lo característico de una distribución teórica es su forma, pero puede presentarse en diferentes posiciones (media aritmética) y de manera más extensa o replegada (desviación tipo). La figura 2 muestra algunos ejemplos diferentes, todos ellos de distribución normal. Observa que siguen siendo campanas.

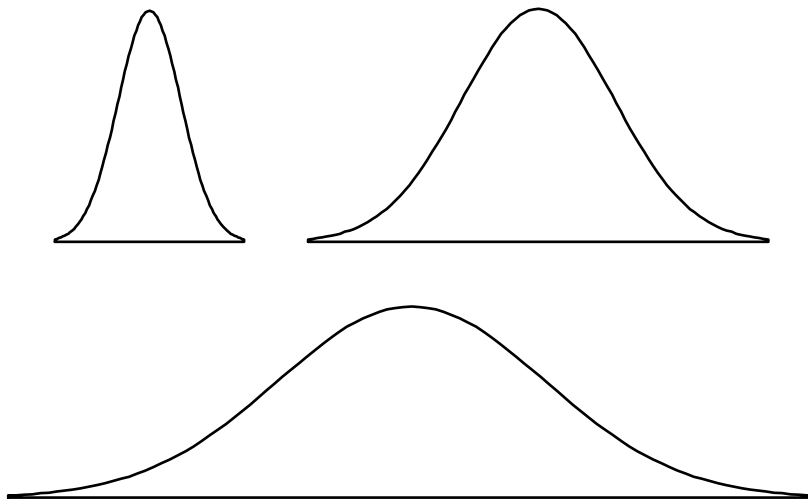


Figura 2: diferentes curvas normales.

Aunque una curva normal puede presentarse con valores muy diferentes de posición y variación, lo habitual es manejar sólo una curva, la estándar o típica, es decir, la que representa distancias estandarizadas (que hemos simbolizado habitualmente con la letra Z). Como recordáis, una puntuación típica está libre de escala de medida (no se mide en metros ni en centímetros) y expresa el número de desviaciones tipo que alejan a

una puntuación directa de la media de su distribución. Por ello, por que las puntuaciones tipo son independientes de la posición y la variación, son las que se utilizan en la distribución normal para ver la relación entre probabilidades y distancias. Vamos a ver algunos ejemplos:

En la curva normal, el 95% de los datos que se encuentran alrededor de la media, se distancian de ésta en 1,96 desviaciones tipo. Es decir, el valor tipificado de los extremos de un intervalo del 95% alrededor de la media es de $Z_{inferior}=-1,96$ y $Z_{superior}=+1,96$. En la figura 3 se muestra ese intervalo.

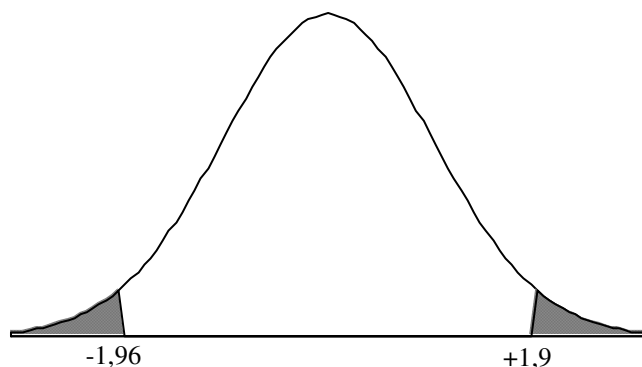


Figura 3: intervalo del 95% en la curva normal.

A dos desviaciones por encima y debajo de la media, se encuentra el 95,5% de los datos, un valor que se utiliza mucho en sociología y que veréis en televisión o en la prensa cuando se publican encuestas. El 99% de los datos, en la curva normal, se encuentran a 2,58 desviaciones tipo (de $Z=-2,58$ a $Z=+2,58$). Estas relaciones entre probabilidad o área, y distancia estandarizada o puntuación típica, son constantes en la curva normal y no importa de qué caso concreto estamos tratando. Es decir, estas relaciones se cumplen para todas las curvas normales de la figura 3 por ejemplo. En otros términos, si una distribución muestral es normal, no importa cuál es su valor esperado o su error tipo, la relación entre Z y área o probabilidad es siempre la misma.

Para suponer que una distribución muestral es normal, se deben cumplir algunos supuestos o condiciones. Vamos a ver dos casos: para la estimación de medias y para la estimación de proporciones.

Se puede suponer que la distribución muestral de medias es normal cuando el tamaño de las muestras es, por lo menos, de 30 unidades (no hemos dejado de hablar del muestreo aleatorio simple. Con otros modelos, las condiciones serían también diferentes). Es decir:

$$n \geq 30$$

Se puede suponer que la distribución muestral de proporciones es normal cuando no sólo el tamaño de las muestras es grande, sino también cuando la proporción no es extrema. Esto es lógico, puesto que conforme la proporción en la población se acerca más a 0 ó a 1, más difícil es que la distribución muestral sea simétrica (y, por lo tanto, normal), puesto que

los resultados muestrales estarán muy cerca del extremo en el que se encuentre la proporción y muy extendidos hacia el otro extremo. Ambos supuestos (n grande y p no extremo) se combinan en dos expresiones que deben cumplirse (ambas):

$$n \cdot p \geq 5 \quad \text{y} \quad n \cdot (1-p) \geq 5$$

Por ejemplo, en el caso de una muestra de $n = 40$, donde se ha obtenido una media aritmética de 12 y una proporción de 0,3 ¿puede suponerse que la distribución muestral es normal, tanto en medias como en proporciones? Veamos:

Como $n = 40 > 30$, se puede suponer que la distribución muestral de medias es normal.

Como $40 \cdot 0,3 = 12 > 5$ y $40 \cdot (1-0,3) = 28 > 5$, entonces se puede suponer que la distribución muestral de proporciones es normal.

Para traducir probabilidades o áreas de la curva normal a distancias estandarizadas, es necesario contar con algún mecanismo. Podría ser una integral definida calculada sobre la función. Pero eso exige mucho esfuerzo. Por suerte hay más procedimientos: recurrir a tablas impresas que relacionan áreas con puntuaciones tipo, o bien acudir al ordenador. Por nuestra parte, esto último es lo que vamos a hacer.

Estimación por intervalo

Ya sabemos qué es una distribución muestral y cuándo suponer que sigue una ley normal. Pero no sabemos qué hacer con ella. Nuestra intención es realizar inferencia y, más concretamente, construir intervalos de confianza. Pero no sabemos qué pinta ahí la distribución muestral más allá del guión que se hizo en el cuadro 1. Bien, pues avanzamos un paso.

Hemos dicho ya que esta inferencia consiste en deducir el valor del parámetro a partir de un estadístico. Vamos a realizar la estimación de una media aritmética. En la muestra hemos obtenido que las personas beben una media de 2,3 litros de alcohol al mes. ¿Cuál es el consumo medio de alcohol al mes en la población? En principio podríamos decir que también 2,3 puesto que no tenemos ninguna otra información. Si hacemos tal cosa y resulta que el consumo es 2,35 ¿hemos acertado? No parece una cantidad muy diferente a la estimada, pero en sentido estricto hemos fallado la estimación, aunque sea muy poco. Lo que hemos hecho se llama *estimación puntual* y consiste en afirmar que el parámetro tiene el mismo valor que el estadístico. En la práctica no hacemos eso, sino que afirmamos que el valor del parámetro es *más o menos* el mismo que el estadístico, porque asumimos cierto error, cierta amplitud en los valores posibles para el parámetro. Así, podríamos afirmar que éste vale $2,3 \pm$

0,2, es decir: de 2,1 a 2,5. Hemos construido una estimación por intervalo.

¿Se puede hacer cualquier intervalo? Pues depende. Está claro que falta algo. No se puede hacer cualquier cosa. Observa los dos siguientes intervalos y piensa en cuál de ellos es más fácil acertar:

1. (2,345 ; 2,346)
2. (2,0 ; 2,6)

En efecto, en el segundo, puesto que contiene más valores. Aquí estamos manejando dos conceptos al mismo tiempo: la amplitud del intervalo y la seguridad para acertar. Y vemos claro que tenemos más seguridad conforme el intervalo es más grande. Pero si pensamos en qué es lo que interesa, llegamos a un problema: lo más útil es construir intervalos precisos (es decir, con un intervalo muy pequeño) con mucha seguridad. Sin embargo, parece que ambos objetivos son incompatibles pues para aumentar la seguridad hay que perder precisión, mientras que para ganar precisión hay que perder seguridad. ¿Cómo hacerlo? En la práctica tenemos sólo dos estrategias: utilizar modelos de muestreo que permitan mucha precisión (aunque esto no está ahora a nuestro alcance) y aumentar el tamaño de la muestra. Cuando la muestra se hace más grande, siendo igualmente aleatoria, entonces los resultados muestrales son más parecidos entre sí y más parecidos al parámetro, por lo que se construyen intervalos más pequeños igualmente seguros. Es decir, se hace más difícil equivocarnos o podemos aspirar a mayor precisión.

Ha llegado el momento de añadir más vocabulario específico:

- Observa cómo hemos construido los intervalos anteriores: al valor del estadístico (2,3) le hemos sumado y restado una misma cantidad. Conforme esa cantidad es más grande, el intervalo se hace también más grande, es decir, más impreciso. Por ello, a esa cantidad que se suma y resta la llamamos *error de precisión*.
- La seguridad la medimos en términos de probabilidad. A la probabilidad de acertar al hacer la estimación (es decir, la probabilidad de que el parámetro se encuentre dentro del intervalo) se le llama *nivel de confianza*, nivel de seguridad o, simplemente, confianza o seguridad. Por el contrario, a la probabilidad de fallar (es decir, de que el parámetro se encuentre realmente fuera del intervalo) se le denomina *riesgo*.

¿Cómo se procede entonces a construir un intervalo? Se sigue el esquema del cuadro 3.

1. Se calcula el estadístico.
2. Se decide un nivel de confianza para hacer la estimación.
3. Utilizando la distribución muestral, se traduce el nivel de confianza en un valor para el error de precisión.
4. Se construye el intervalo.

Cuadro 3: proceso en la construcción de un intervalo de confianza.

El cálculo del estadístico no compete a este documento. Se entiende que ya sabemos obtener medias, proporciones o lo que proceda en cada caso. Lo que vamos a hacer en los siguientes apartados es resolver los tres puntos restantes del cuadro 3, desde el nivel de confianza a la construcción del intervalo.

Cómo decidir un nivel de confianza

El nivel de confianza, tal y como lo hemos definido, es la probabilidad de acertar al realizar la estimación. Su complementario es el riesgo o probabilidad de fallar al realizar la estimación. Por ello, es importante pensar en cuáles son las consecuencias que se derivan de errar y decidir la probabilidad en función de ello.

Piensa por un momento que te invitan a saltar desde un avión, en paracaídas. Aceptas. Subes. Te tiras. Haces unas piruetas en el aire con el resto de las personas que participan contigo en esta aventura. En un momento, se decide tirar de la cinta para que se accione el paracaídas. Buscas la cinta. La tienes ya en la mano. Lees lo que pone: "probabilidad de que el mecanismo falle cuando estire de la cinta: 5%". ¿Qué te parece? Más o menos, este lema significa que, de cada 100 veces que tires de la cinta, en cinco ocasiones no se accionará el paracaídas. Aunque suponemos que no tendrás oportunidad de realizar cinco saltos con errores. ¿Qué probabilidad de error te parece ya asumible? ¿Una de cada cien veces (1%)? ¿Una de cada mil (0,1%)? ¿Una de cada millón (0,0001%)? Los paracaídas son mecanismos generados por las personas, es decir, imperfectos. Es imposible que no fallen nunca. No pensar en ello no elimina la probabilidad de error.

Piensa ahora que cada piedra de un camino, que alcance determinado tamaño, contiene un valor numérico pintado en la superficie: es la probabilidad de que bajo ella no se encuentre oro. Si le das una patada a una piedra o si te agachas para darle la vuelta y ver qué hay debajo, implicas un esfuerzo que puede no servir para nada o puede que resulte recompensado con una cantidad variable de oro. ¿Cuál es la probabilidad de errar al suponer que hay oro bajo la piedra? Eso es lo que pone en la superficie. Pero ¿qué cantidad es la que marca el límite entre estar dispuesto a realizar el esfuerzo de levantarla o no? ¿Un 5%? Ello

implicaría que de cada 100 intentos, en 5 no habría oro. Posiblemente, estaríamos dispuestos a correr riesgos superiores al 99% (de cada cien intentos, en 99 no hay oro).

Como vemos, las consecuencias de errar en la estimación son muy diferentes en los dos ejemplos anteriores. Ello exigiría que, atendiendo a tales consecuencias, definiéramos un riesgo o un nivel de confianza específico y adecuado. Pero en la práctica esto no se hace, por tres motivos:

1. Fisher reflexionó sobre ello hace casi un siglo y propuso la cantidad del 5% como riesgo, es decir, del 95% como nivel de confianza. Surgió de unos cálculos sobre la probabilidad que exigía a una prueba donde una mujer afirmaba que era capaz de discriminar si en una taza de té con leche se había echado antes el té o la leche. Como Fisher pensó en ello y era una persona de amplio prestigio (sigue siéndolo), es difícil discutirle, impone.
2. Pensar es laborioso. Se ha instaurado ya como un hábito escoger el riesgo del 5% en situaciones normales o del 1% cuando se quiere ser más exigente. Es tan habitual que se da por asumido y ello permite que nos relajemos a la hora de enfrentarnos a la ardua tarea de cuantificar unas consecuencias en términos finales de probabilidad.
3. Dado lo extenso de esta práctica, cuando a alguien se le ocurre escoger un nivel de confianza que no es del 95% ni del 99%, lo normal es que se le pidan explicaciones. Si no quieres darlas, haz lo mismo que los demás. Este consejo que parece sacado de un libro barato de costumbres o de moral a la medida de la conveniencia, es lo que se hace en la práctica de la investigación.

Cómo se traduce un nivel de confianza en un valor para el error de precisión

El error de precisión ha sido definido aquí como el radio del intervalo de estimación, es decir, como la cantidad que se le resta y se le suma al estadístico para construir el intervalo en el que esperamos que se encuentre el parámetro. Pero hay otra definición que nos permite no sólo llegar mejor al concepto de error de precisión, sino que ayuda a traducir el nivel de confianza a partir de la distribución muestral. Y eso es lo que vamos a hacer ahora.

Decir "espero que el parámetro se encuentre en este intervalo alrededor del estadístico", es lo mismo que decir "espero que la distancia entre el parámetro y el estadístico no sea mayor que el radio de este intervalo". En otros términos: el error de precisión es la máxima distancia que cabe esperar entre el estadístico y el parámetro. Esperamos que eso sea así con una probabilidad que hemos llamado nivel de confianza. Luego, lo que haremos para hacer la traducción es calcular la distancia que se

corresponde al nivel de confianza en la distribución muestral (la distribución de lo que cabe esperar que ocurra).

Observa la figura 4. En ella hemos representado un intervalo del 95% en una distribución muestral de medias que sigue una ley normal.

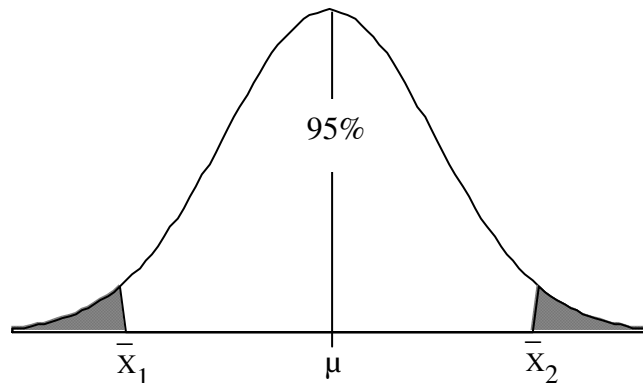


Figura 4: 95% de resultados muestrales.

Observa que hemos representado el valor esperado de la media aritmética con μ , que es el símbolo que representa la media en la población. Esto es así porque realmente coinciden, es decir, la media de la distribución muestral de medias coincide con la media de la población (no es una feliz coincidencia, se puede demostrar matemáticamente, pero estoy seguro que entenderéis que no nos ocupamos de ello aquí). Del mismo modo, la media de la distribución muestral de proporciones (valor esperado de la proporción) coincide también con la proporción en la población, es decir, π .

En la representación gráfica de la figura 4 se ha representado un intervalo. Es el que contiene al 95% de todos los resultados muestrales posibles, es decir, a los valores de la media aritmética más cercanos a μ que provienen del 95% de las muestras que se extraen de esa población (con un tamaño concreto y de un modo concreto). Desconocemos cuál es el valor de μ , por lo que también desconocemos cuáles son los valores de las dos medias que se encuentran en los límites del intervalo. Pero sí que podemos calcular cuál es la distancia que las separa. Y podemos calcularlo porque sabemos que la distribución es normal y ya hemos visto que ello nos permite traducir áreas o probabilidades a distancias.

En puntuaciones tipo o distancias estandarizadas, sabemos que un área del 95% central en la curva normal se corresponde con una $Z = \pm 1,96$. Como una distancia estandarizada es el número de desviaciones tipo que separa una puntuación con respecto a la media de su distribución, entonces, esa Z es el número de errores tipo que separan el extremo con respecto a μ . Por ejemplo, pensemos en un intervalo del 95%, en el que $Z = \pm 1,96$, que es igual a la distancia expresada en número de desviaciones tipo (en la distribución muestral, número de errores tipo); y ara un error tipo de valor 0,84. La distancia que estamos calculando es la máxima que cabe esperar en el 95% de los casos, es decir, lo que hemos

denominado error de precisión. Si simbolizamos el error tipo con et , y el error de precisión con ep , entonces:

$$Z = \frac{\text{distancia}}{\text{desviación}} = \frac{ep}{et} \Rightarrow ep = Z \cdot et = 1,96 \cdot 0,84 = 1,65$$

Observa que hemos traducido un nivel de confianza en un error de precisión mediante dos cálculos intermedios: de nivel a distancia estandarizada y de ésta a error de precisión. Para hacer estos cálculos necesitamos fórmulas, tablas o el ordenador. Vamos a utilizar este último. De hecho, veremos algunos ejemplos nada más terminando esta explicación que, tal vez, esté resultando ser demasiado larga.

Cómo construir el intervalo de confianza

El error de precisión es la máxima distancia que cabe esperar que separe al estadístico del parámetro. Luegz, al sumar y restar el error de precisión al estadístico, construimos un intervalo alrededor de éste donde esperamos que se encuentre el parámetro, puesto que no puede estar más lejos que esa máxima distancia. Para ser más correctos explicando este proceso, no decimos que "no puede estar más lejos", sino que "esperamos que no esté más lejos", ya que se trata de una máxima distancia *esperada*. Que lo esperemos no significa que ocurra así exactamente. Ya sabemos que la probabilidad de que ocurra así es 95% (en el ejemplo).

Supongamos que el valor en la muestra es $\bar{X} = 9$. En tal caso, la estimación por intervalo genera:

$$\mu \in \{9 \pm 1,65\}_{0,95} = \{7,35; 10,65\}_{0,95}$$

Lo anterior es una expresión que permite indicar de qué intervalo estamos hablando. Se lee así: con un nivel de confianza del 95% estimamos que la media aritmética de la población se encuentra entre los valores 7,35 y 10,65.

Vamos a ver cómo realizamos esta gestión con SPSS. Para ello, hay que distinguir entre la estimación de medias aritméticas y la estimación de proporciones.

Estimación de medias con SPSS

- Paso 1: abrir un archivo de datos.

Vamos a utilizar un archivo de datos real, que obtuvimos desde la clase¹. En el anexo se encuentran las variables y los datos. Los 84 estudiantes generaron preguntas y seleccionaron después seis de ellas, que se aplicaron en forma de cuestionario. Con ese archivo de datos vamos a realizar las estimaciones, tanto de medias como de proporciones.

- Paso 2: identificar la variable con la que realizamos el análisis.

La variable va a ser "¿Eres feliz?" (feliz).

- Paso 3: ejecutar el comando Explorar.

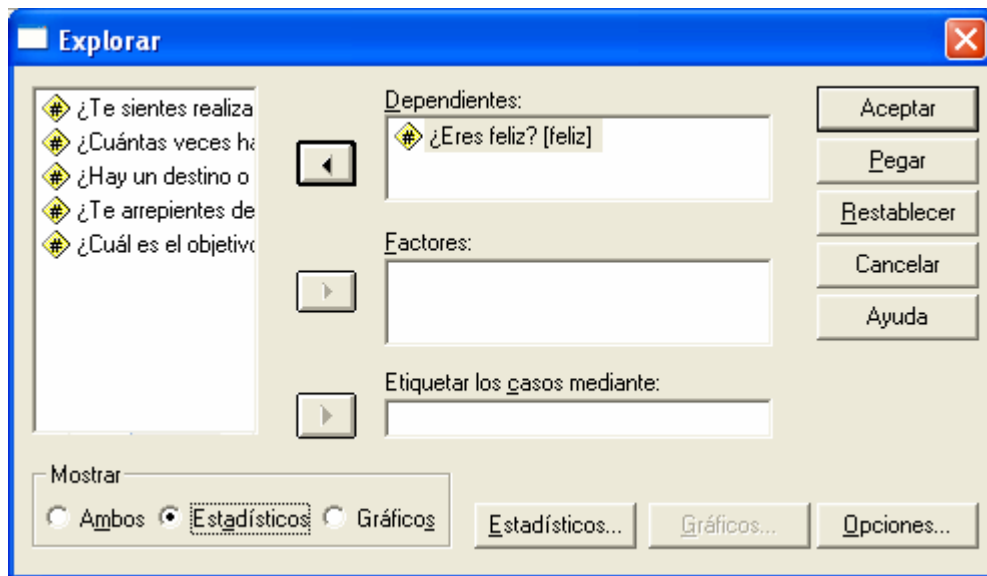


Figura 5: cuadro de diálogo de Explorar.

Con "Análizar / Estadísticos descriptivos / Explorar..." se abre un cuadro de diálogo desde el que obtendremos el intervalo. Para ello, pulsa en el nombre de la variable y llévala al recuadro "dependientes". Sólo queremos estadísticos, así que desactiva "Mostrar *ambos*" activando "Mostrar *Estadísticos*". La figura 5 muestra la apariencia que debe tener el cuadro de diálogo después de realizar estas gestiones. Al pulsar sobre el

¹ Las estimaciones estadísticas tienen sentido porque provienen de muestras aleatorias, ya lo hemos dicho. No estamos respetando, no obstante, este principio. Hemos acudido al archivo proveniente de los datos de una clase de la universidad. No hubo realmente una selección aleatoria desde una población. Lo que hacemos es un ejercicio de imaginación. Contamos con datos que nos interesan y que resultan atractivos porque provienen del mismo grupo de clase. Imaginemos que se corresponden realmente con una muestra aleatoria de una población concreta, como pueden ser los estudiantes universitarios en general. Dado que este ejemplo no vamos a utilizarlo fuera de este documento, es decir, dado que no se va a aplicar para tomar decisiones sobre la sociedad, sabemos que no generamos daño con ello. Pero no hay que abandonar el principio de que la inferencia estadística sólo tiene sentido si las muestras han sido aleatorias.

botón “Estadísticos”, se abre otro cuadro de diálogo (obsérvalo en la figura 6). En él podemos controlar el nivel de confianza que, por defecto, se encuentra al 95%.

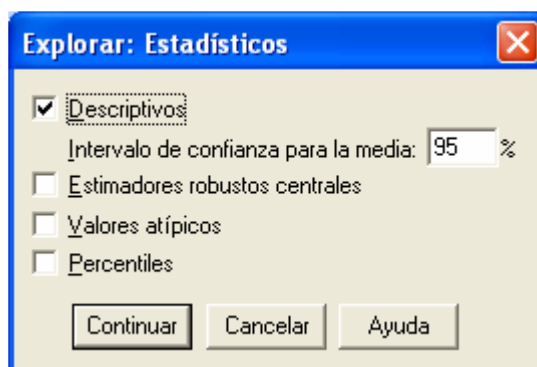


Figura 6: cuadro de diálogo para la especificación del nivel de confianza.

Así lo dejamos, de momento. Pulsa en “Continuar” para volver al cuadro de diálogo de la figura 5 y ahí pulsa “Aceptar”. Obtendrás como salida un listado amplio de estadísticos, que posiblemente ya conozcas por la realización de estudios descriptivos previos. La primera información que aparece es el valor de la media aritmética (7,21) y los extremos inferior (6,84) y superior (7,58) del intervalo de confianza del 95%.

De esta forma, el intervalo queda así:

$$\mu \in \{6,84; 7,58\}_{0,95}$$

En otros términos, con un nivel de confianza del 95% afirmamos que la media de felicidad manifiesta en la población (en una escala entre 0 y 10) se encuentra entre 6,84 y 7,58.

Ensayá diferentes variables e intervalos de confianza. En las dos siguientes tablas (tablas 1 y 2) tienes algunos ejemplos. En el primer caso, para un nivel de confianza del 95%. En el segundo caso, para el 99%:

Variable	Intervalo del 95%	
	desde...	hasta...
¿Eres feliz?	6,84	7,58
¿Te sientes realizado/a?	6,06	6,63
¿Cuántas veces has sido infiel?	0,49	1,10

Tabla 1: intervalos del 95% para algunas variables.

Observa cómo, al aumentar el intervalo de confianza, es decir, al exigir al procedimiento mayor seguridad en la estimación, aumenta también el error de precisión y, con él, la amplitud del intervalo. La segunda tabla, con respecto a la primera, muestra intervalos más seguros, pero menos precisos.

Variable	Intervalo del 99%	
	desde...	hasta...
¿Eres feliz?	6,62	7,70
¿Te sientes realizado/a?	5,97	6,72
¿Cuántas veces has sido infiel?	0,39	1,20

Tabla 2: intervalos del 99% para algunas variables.

Estimación de proporciones con SPSS

SPSS no cuenta con un procedimiento similar al que hemos visto en el subapartado anterior para el caso de la estimación de proporciones. Para conseguirlo, hay que utilizar una pequeña trampa. Ocurre que una proporción puede ser considerada un caso particular de media aritmética si se codifica convenientemente.

Imagina que hemos medido el grado de felicidad en tres categorías: triste, normal y feliz. Queremos calcular la proporción de personas que se manifiestan felices. Para ello, codificamos la categoría que nos interesa (*feliz*) con un 1, mientras que el resto (*triste* y *normal*) se codifican con un 0. En el cuadro 3 constan los datos originales y su codificación.

triste	0	normal	0	normal	0
triste	0	triste	0	feliz	1
normal	0	feliz	1	feliz	1
feliz	1	normal	0	feliz	1
feliz	1	feliz	1	triste	0
feliz	1	triste	0	feliz	1
feliz	1	normal	0	normal	0
normal	0	feliz	1	feliz	1
feliz	1	normal	0	normal	0
triste	0	feliz	1	feliz	1

Cuadro 3: codificación 0-1 para la variable *sensación de felicidad*.

Si cuentas cuántos datos hay y cuántos se corresponden con el valor *feliz*, encontrarás, respectivamente, 30 y 15. Hablamos, pues de un 50% de datos con el valor *feliz* (una proporción de valor 0,5). Ahora calcula la media aritmética de los 30 datos 0-1. Has encontrado lo mismo ¿verdad? No es un misterio. Si bien se puede demostrar matemáticamente, observa que las operaciones que has realizado son las mismas: al sumar todos los datos es como si hubieras contado el número de ellos que tienen el valor 1, puesto que los datos con valor 0 no añaden nada a la suma. Observa, además, que la media y la proporción se obtienen dividiendo por el mismo número: el total de datos.

Así pues, la proporción es un caso particular de media aritmética cuando se codifica de tal forma que la categoría cuya proporción queremos

calcular se codifica con el valor 1, mientras que el resto de las categorías se codifican con 0. Gracias a ello, lo que hacemos en SPSS es codificar convenientemente y pedir una estimación de media aritmética con el comando Explorar, tal y como hemos visto en el subapartado anterior.

- Paso 1: abrir un archivo de datos.

Vamos a utilizar de nuevo el mismo archivo que para las medias.

- Paso 2: identificar la variable con la que realizamos el análisis.

La variable va a ser "¿Te arrepientes de haber nacido?" (nacido). Lo que nos interesa de ella es estimar la proporción de personas en la población que dirían en la encuesta que se arrepienten de haber nacido. Si pides una tabla de frecuencias (Analizar / Estadísticos descriptivos / Frecuencias...) Obtendrás el resultado de la tabla 3.

¿Te arrepientes de haber nacido?

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos no	81	96,4	96,4	96,4
sí	3	3,6	3,6	100,0
Total	84	100,0	100,0	

Tabla 3: tabla de frecuencias para el arrepentimiento por haber nacido.

La proporción que le corresponde a la categoría cuya proporción en la población queremos estimar es de 3,6%. Como esta variable ya se encuentra codificada de manera conveniente (el valor que nos interesa, es decir, la respuesta *sí* se codifica con 1, mientras que *no* está codificada con 0), podemos proceder directamente a la petición del intervalo de confianza.

- Paso 3: ejecutar el comando Explorar.

Al realizar este paso, de forma idéntica al caso de las medias aritméticas del subapartado anterior, obtenemos que

$$p \in \{0,00; 0,08\}_{0,95}$$

Es decir, cabe esperar (con un nivel de confianza del 95%) que entre un 0% y un 8% de la población se arrepienta de haber nacido.

Pero pongámonos ahora en el caso de que la codificación no es la conveniente, en cuyo caso hay que recodificar las variables para adaptarlas a las condiciones de la estimación. Para verlo, nos planteamos estimar la proporción de personas que, en la población, tienen como principal objetivo ser felices. Al pedir a SPSS una tabla de frecuencias de la variable sobre el objetivo de la vida, el resultado es:

¿Cuál es el objetivo de tu vida?

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	terminar la carrera	4	4,8	4,9	4,9
	ser feliz	54	64,3	65,9	70,7
	ser libre	8	9,5	9,8	80,5
	ser independiente	7	8,3	8,5	89,0
	vivir dignamente	6	7,1	7,3	96,3
	formar una familia	3	3,6	3,7	100,0
	Total	82	97,6	100,0	
Perdidos	Sistema	2	2,4		
Total		84	100,0		

Tabla 4: tabla de frecuencias para la variable *¿Cuál es el objetivo de tu vida?*

“Ser feliz” ha sido la respuesta escogida por 54 de las 84 personas de la muestra, es decir, por un 64,3%. Hemos de recodificar esta variable de tal modo que el valor 1 se corresponda con el valor “Ser feliz”, mientras que el resto de las categorías se recodifiquen con el valor 0. Para ello, hay que acudir a “Transformar / Recodificar / En distintas variables...”. Hemos escogido la opción de recodificar en una variable distinta a la original para mantener ésta para futuros análisis. La nueva variable se puede denominar “profeliz”. Para ello, hay que realizar los siguientes pasos:

1. Pulsa sobre la variable “objetivo” y llévala al recuadro “Variable de entrada”.
2. En el recuadro “Variable de resultado / Nombre:” escribe “profeliz”.
3. Pulsa sobre el botón “Cambiar” para que se actualice esta información. La apariencia de este cuadro de diálogo se muestra en la figura 7.

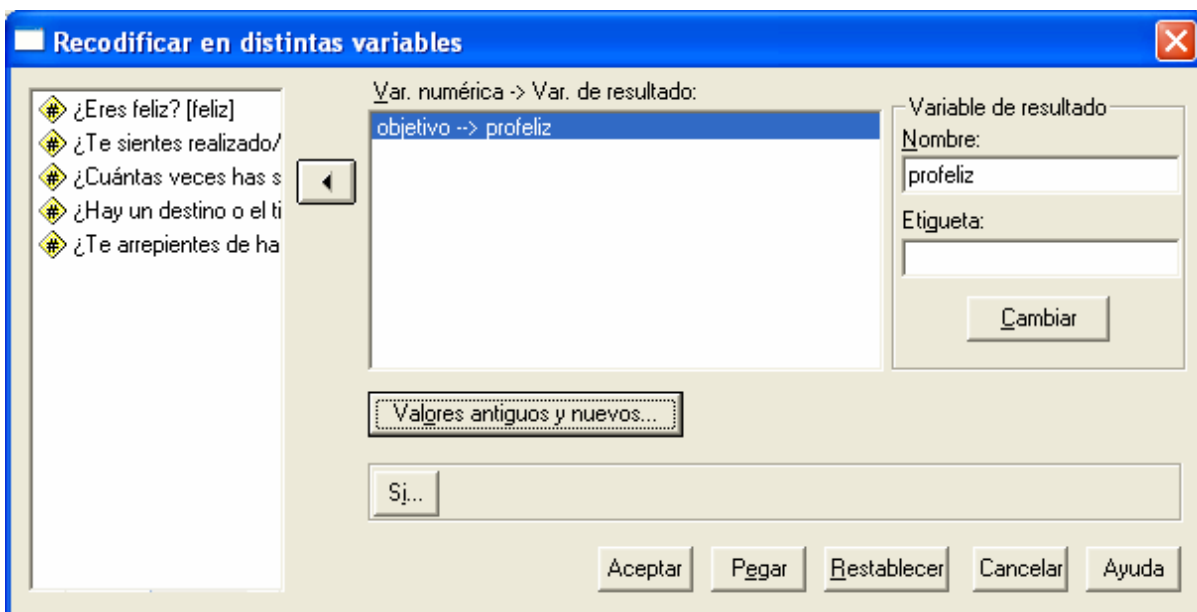


Figura 7: cuadro de diálogo de recodificar en distintas variables.

4. Pulsa sobre el botón “Valores antiguos y nuevos...” para introducir los cambios.

5. En el nuevo cuadro de diálogo, introduce el valor "2" en el recuadro "Valor antiguo / valor:", puesto que 2 es el valor que tiene la categoría "ser feliz" en la variable "objetivo".
6. En el recuadro "Valor nuevo / valor:", introduce el valor "1".
7. Pulsa el botón "Añadir" para que se registre esta recodificación.
8. Activa "Valor antiguo / Todos los demás valores".
9. En el recuadro "Valor nuevo / valor:", introduce el valor "0".
10. Pulsa de nuevo "Añadir" y ya hemos terminado de especificar los cambios (la apariencia del cuadro de diálogo se encuentra en la figura 8).
11. Pulsa sobre el botón "Continuar" para acceder al cuadro de diálogo principal del procedimiento (figura 7).
12. Y, por último, pulsa en "Aceptar".

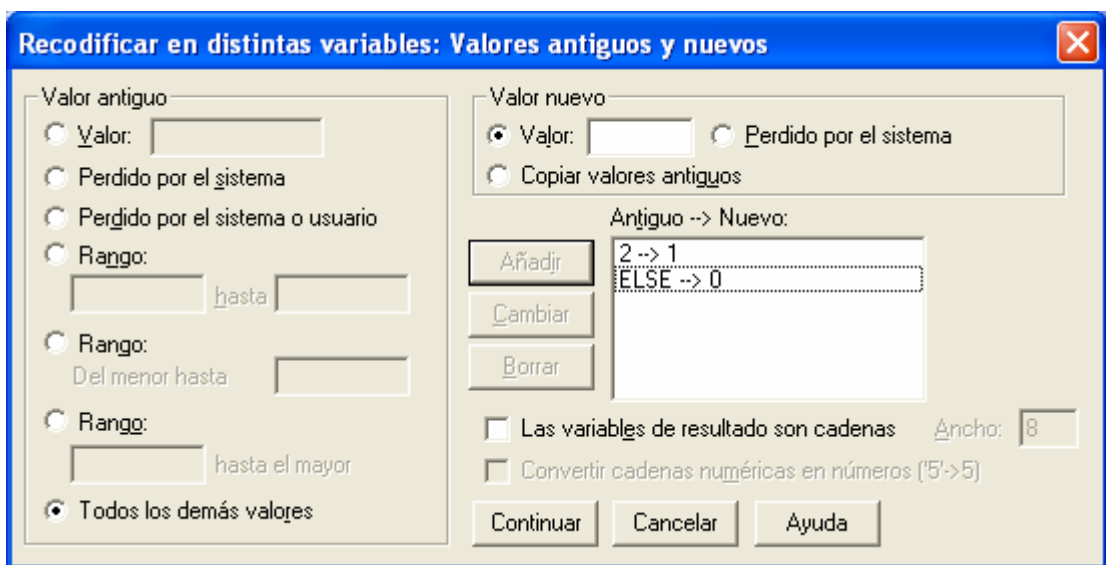


Figura 8: cuadro de diálogo para la especificación de los cambios.

Si pides ahora a SPSS una tabla de frecuencias de la nueva variable (profeliz), el resultado es el que consta en la tabla 5.

profeliz

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
,00	30	35,7	35,7	35,7
1,00	54	64,3	64,3	100,0
Total	84	100,0	100,0	

Tabla 5: tabla de frecuencias para *profeliz*.

Con esta variable, actuamos como con el caso de *nacido*. Al recurrir a Explorar y pedirle los descriptivos con el intervalo de confianza del 95%, obtenemos que

$$n \in \{0,538; 0,748\}_{0,95}$$

Es decir, que entre un 53,8% y un 74,8% de la población tienen como principal objetivo de su vida ser felices.

Ensayo este procedimiento con la variable *infidel*. Recodifícala para crear una nueva variable donde sólo existan dos categorías: nunca ha sido infiel (valor 0) y lo ha sido al menos una vez (valor 1). En tal caso, la estimación por intervalo con un nivel de confianza del 95% será:

$$n \in \{0,198; 0,397\}_{0,95}$$

Pruebas de significación de la hipótesis nula

Como, sospecho, aún no has olvidado, las pruebas de significación de la hipótesis nula es el otro campo que no hemos abordado todavía, de los dos que nos ocupan. Toca ahora.

El objetivo de una prueba de significación de la hipótesis nula es facilitar la toma de decisiones con respecto a la existencia de hipótesis de investigación que permiten ser respondidas con un sí o un no. Algunos ejemplos son ¿Ha cambiado la actitud frente al machismo? ¿Existe relación entre la depresión y la edad? ¿Son los europeos más sensibles a la música?

El recurso lógico que se utiliza en una prueba de significación de la hipótesis nula se parece al del sistema de justicia:

1. Se lleva a alguien a juicio porque se sospecha que ha cometido un delito.
2. El proceso comienza con una premisa inicial: esa persona es inocente mientras no se demuestre lo contrario.
3. Se presentan evidencias o hechos que permiten plantearse la credibilidad de la premisa inicial.
4. A la luz de lo que deduce de esas evidencias o hechos, se toma una decisión: ¿se mantiene o se rechaza la hipótesis de inocencia?
5. En función de la decisión, tiene lugar una conclusión: si se rechaza la hipótesis de inocencia, se declara culpable al acusado. En caso contrario, el juicio termina dejándolo libre.

En una prueba de significación de la hipótesis nula (PSHN) se lleva un proceso similar, que puede ser expuesto con los mismos puntos (de una forma más esquemática, el proceso se expone en el cuadro 4):

1. Se sospecha que existe alguna relación entre variables o alguna diferencia de algún tipo.
2. El proceso comienza con una hipótesis inicial (llamada hipótesis nula y representada con el símbolo H_0): no existe relación entre las variables o no hay diferencia entre los términos que se comparan, mientras no se demuestre lo contrario.
3. Se presentan evidencias o hechos que permiten plantearse la credibilidad de la H_0 .

4. A la luz de lo que deduce de esas evidencias o hechos, se toma una decisión: ¿se mantiene o se rechaza la H_0 ?
5. En función de la decisión, tiene lugar una conclusión que traduce el rechazo o el mantenimiento de la H_0 con los mismos términos con que se enunció el problema de investigación.

- 1) Enunciado de H_0 .
- 2) Estudio empírico.
- 3) Decisión.
- 4) Conclusión.

Cuadro 4: esquema de una PSHN.

Para explicar este proceso, vamos a seguir un caso concreto. A partir de nuestro archivo de datos vamos a dar respuesta a un par de interrogantes:

- (1) Se tiene la creencia de que las personas de la población tienen un nivel de felicidad medio (valor aproximadamente de 5 en una escala de 0 a 10). ¿Puede considerarse correcta esa creencia?
- (2) ¿Existe relación entre el grado en que alguien se siente feliz y el grado en que se encuentra realizado o realizada en su vida?

Hay varios procedimientos para abordar el paso 2, pero aquí sólo nos interesan dos: el que parte del intervalo de estimación (que ya conocemos) y el que se basa en el cálculo de una probabilidad que representa la *significación estadística*. Estos procedimientos se van a mostrar con los dos interrogantes expuestos.

PSHN mediante un intervalo de confianza

El interrogante al que intentamos dar respuesta es:

Se tiene la creencia de que las personas de la población tienen un nivel de felicidad medio (valor aproximadamente de 5 en una escala de 0 a 10). ¿Puede considerarse correcta esa creencia?

El primer paso es enunciar la hipótesis nula: "No existe diferencia entre el grado de felicidad medio que midamos ahora y 5". En otros términos: "El grado de felicidad medio en la población es 5".

El estudio empírico consistirá en obtener una muestra aleatoria simple, calcular los estadísticos necesarios y observar cuán creíble resulta la H_0 a la luz de esos resultados. La muestra ya la tenemos y ya sabemos también cómo calcular un intervalo de confianza. Lo hacemos y obtenemos el siguiente resultado:

$$\mu \in \{6,84; 7,58\}_{0,95}$$

Es decir, con un 95% de confianza, creemos que el grado medio de felicidad de la población se encuentra entre 6,84 y 7,58. Observa que eso es lo que creemos. ¿Se encuentra ahí el valor defendido por la hipótesis nula, es decir, 5?

Ahora estamos en disposición de abordar el tercer paso: tomar una decisión. Recuperemos la pregunta clave en este procedimiento: ¿Se encuentra en el intervalo de estimación el valor defendido por la hipótesis nula, es decir, 5? Pues no, 5 no está en el intervalo, es decir, no es uno de los valores que esperamos para el grado medio de felicidad en la población. Por tanto, ya sabemos qué hacer con la hipótesis nula: rechazarla. No la mantenemos porque los datos no son compatibles con ella. En otros términos, a la luz del estudio empírico no es creíble la hipótesis de que el valor del parámetro sea 5. El términos ortodoxos, la decisión se podría redactar como sigue: *con un riesgo de error del 5%, se rechaza la hipótesis nula.*

Por último, ¿cuál debe ser, por tanto, la conclusión? Recordemos la pregunta: *¿Puede considerarse correcta la creencia de que el grado medio de satisfacción en la población es 5?* La conclusión es, pues, "No, no puede considerarse tal cosa, el grado medio de satisfacción en la población no es 5".

PSHN mediante significación estadística

El interrogante al que intentamos dar respuesta es:

¿Existe relación entre el grado en que alguien se siente feliz y el grado en que se encuentra realizado o realizada en su vida?

Como índice de relación, vamos a recurrir a la rho de Spearman que, como sabes, se trata de una prueba no paramétrica de relación entre dos variables cuantitativas u ordinales. Pero, como ya se ha insistido suficientemente en otro momento, no podemos abordar directamente una prueba de relación sin antes realizar un estudio descriptivo. En este caso, generamos un diagrama de dispersión entre ambas variables (Gráficos / Dispersión / Simple). El resultado se encuentra en la figura 9.

Observa que la relación es aproximadamente lineal, si bien hay dos puntos que se alejan sensiblemente de esa tendencia. Uno de los puntos tiene las coordenadas (3,9), es decir: un 3 en *¿Te sientes realizado/a?* y un 9 en *¿Eres feliz?* El otro punto tiene las coordenadas (8,1). Son dos casos que distorsionan la tendencia general (son los casos 26 y 2, respectivamente). Lo que procede aquí es eliminarlos del análisis. Hay muchas posibilidades para ello. La más sencilla es borrar el valor de estos dos casos en una de las dos variables (en *¿Eres feliz?*, por ejemplo) y volver a escribirlo tras el análisis. También podríamos crear

automáticamente una nueva variable que contenga los mismos valores que *¿Eres feliz?* pero donde mostrara un valor perdido en los casos implicados. Sea como fuere, procedemos de tal modo y calculamos después el índice rho de Spearman.

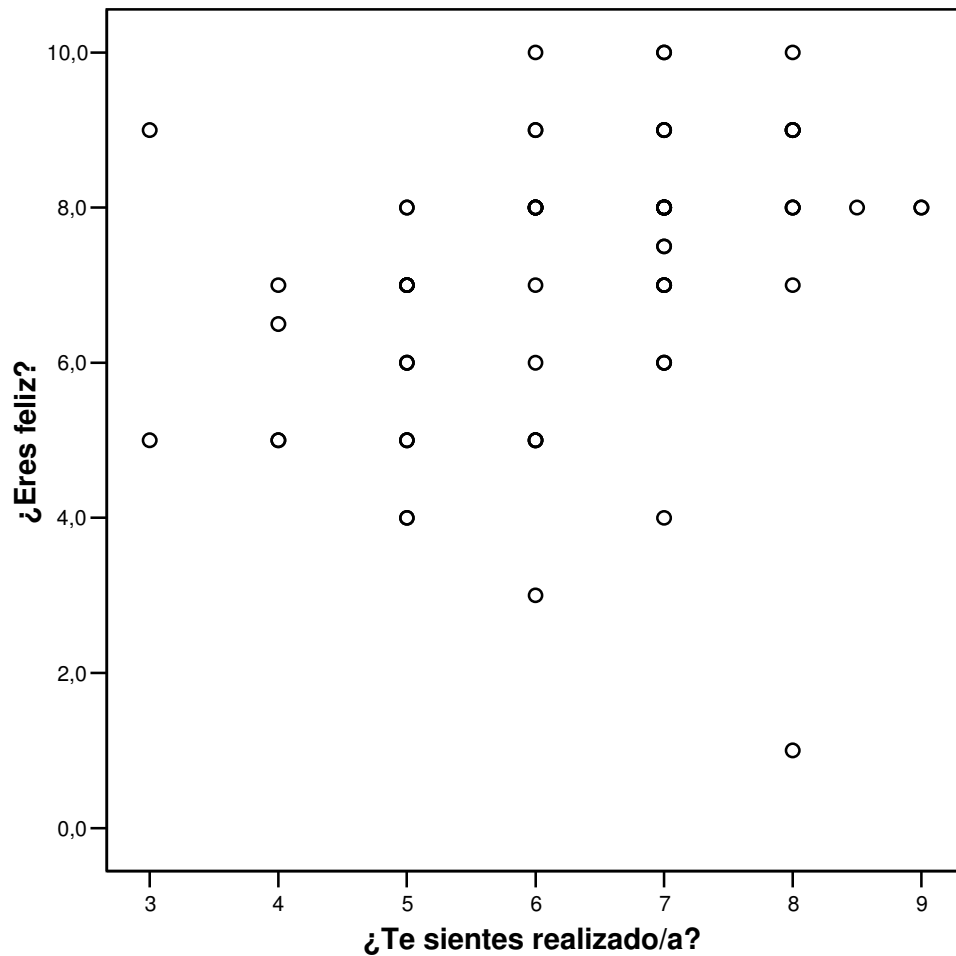


Figura 9: gráfico de dispersión entre dos variables ordinales.

Correlaciones

			¿Eres feliz?	¿Te sientes realizado/a?
Rho de Spearman	¿Eres feliz?	Coefficiente de correlación	1,000	,566(**)
		Sig. (bilateral)	.	,000
		N	81	81
	¿Te sientes realizado/a?	Coefficiente de correlación	,566(**)	1,000
		Sig. (bilateral)	,000	.
		N	81	84

** La correlación es significativa al nivel 0,01 (bilateral).

Tabla 6: resultado al pedir la rho de Spearman en SPSS.

Para calcular el índice, se acude a

Analizar / Correlaciones / Bivariadas...

Activa, a continuación, la casilla de verificación de Spearman y desactiva la casilla de Pearson, que aparece activa por defecto. Selecciona las dos variables que interesan y pulsa el botón "Aceptar". El resultado es el que se muestra en la tabla 6.

Como sabemos, el resultado de la correlación va de -1 a +1. El primero expresa máxima relación negativa (conforme una de las dos variables aumenta de valor, la otra disminuye). El segundo se refiere a una máxima relación positiva (conforme una de las variables aumenta de valor, la otra también lo hace). El resultado de correlación que se muestra en la tabla 6 (0,566) es positivo y de cuantía importante. Así que parece que, al menos en el contexto de los datos de la muestra, las dos variables están relacionadas. Pero, siguiendo con la inquietud general de la inferencia, ¿se puede considerar del mismo modo en la población? Para responder a esta pregunta, acudimos a una PSHN.

El primer paso es enunciar la hipótesis nula. Ésta va a establecer que no existe relación entre ambas variables, que el valor de la correlación en la población es 0. Lo que ocurre es que hay fluctuaciones aleatorias entre muestras y, por ello, hemos podido encontrar casualmente algo diferente de 0. Pero esa diferencia no será importante.

La forma de abordar la credibilidad de esa hipótesis va a consistir, ahora, en cuantificar, en términos de probabilidad, la distancia entre lo encontrado en la práctica y lo esperado según la hipótesis nula. Para ello, retomamos la distribución muestral, representando en ella lo que hemos encontrado. La distribución muestral que utilizamos es la que supone cierta la hipótesis nula, es decir, es la que parte de que el valor de la correlación en la población es exactamente 0. Y en ella, representamos lo que se ha encontrado: 0,566. La figura 10 muestra esa representación.

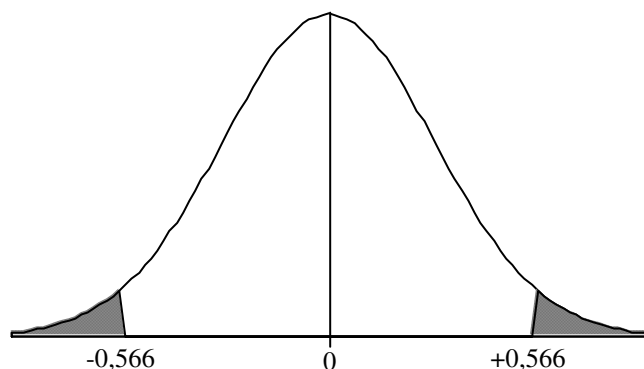


Figura 10: medida de la significación de un resultado muestral.

Observa que hemos rayado dos zonas. El resultado obtenido es +0,566 y lo que pretendemos es calcular cuán significativo es, es decir, cuán

importante es la distancia que separa 0,566 del valor esperado según la hipótesis nula (0). Para ello, razonamos del siguiente modo:

- Lo que interesa es señalar los valores tanto o más raros como el encontrado, puesto que son los valores que nos llevan a dudar de la hipótesis nula (porque se encuentran muy alejados).
- Tan raro es el valor 0,566 como -0,566, puesto que ambos ponen en cuestión del mismo modo el valor esperado (0).
- Y más raros son aún los que se encuentran más alejados de la hipótesis nula, es decir, los menores de -0,566 y los mayores de +0,566.
- El área rayada representa la probabilidad de encontrar todos esos resultados muestrales. Son los que llegan a ser tanto o más raros que lo encontrado en la muestra.

¿Por qué nos interesa medir la probabilidad de encontrar esos valores raros? Porque constituye un riesgo fundamental: el de equivocarnos si decimos que la muestra donde encontramos esos valores no proviene de la población que representa la hipótesis nula. En otros términos: vemos que esos resultados muestrales se encuentran presentes en la distribución muestral que se deriva de la hipótesis nula. Si la rechazamos, estamos corriendo un riesgo, ¿cuál? El riesgo es la probabilidad de que realmente esos resultados sean posibles suponiendo cierta la hipótesis nula. Es lo que hemos representado en la figura 10. Por ejemplo, si el área rayada es un 3% del total de la distribución muestral, significa dos cosas: (1) que 0,03 es la probabilidad de encontrar resultados tanto o más extremos que 0,566 provenientes de una población donde el valor es 0; y (2) que, por tanto, 0,03 es también la probabilidad de equivocarnos si afirmamos que el valor en la población no es 0.

Ese riesgo (el de equivocarnos al rechazar la hipótesis nula de que $\rho = 0$) se denomina también "significación estadística", "grado de significación" o, sencillamente "significación". El nombre tiene sentido, puesto que mide en qué grado lo encontrado tiene significado en términos estadísticos.

En nuestro caso, SPSS calcula la significación (lo hace prácticamente en todas las pruebas de relación) y nos dice que es 0,000 (se encuentra en la tabla 6. Así que 0,000 es el riesgo que corremos de equivocarnos al rechazar la hipótesis nula. La pregunta ahora es ¿corremos ese riesgo? ¿Rechazamos la hipótesis nula sabiendo que nos podemos equivocar con una probabilidad de 0,000? La respuesta aquí es obvia: pues claro que corremos ese riesgo, porque su valor es 0, es decir, no hay riesgo. En otros términos, podemos decir que tenemos una plena seguridad de que la correlación, en la población, no es cero. Unido esto a nuestra conclusión a nivel de la muestra, concluiremos que también a nivel de la población, las variables *¿Eres feliz?* y *¿Te sientes realizado/a?* están relacionadas entre sí.

El ejemplo no es muy bueno en esto último, puesto que en la práctica es difícil encontrar una significación tan pequeña que, redondeada a tres cifras decimales, sea 0. ¿Qué hacer cuando el riesgo es mayor que 0? Respuesta: depende de cuán mayor sea. En otros términos: cuando la significación supera un límite, decimos que el riesgo es excesivo y que no vamos a asumirlo, es decir, que no vamos a rechazar la hipótesis nula, es decir, que la mantenemos. ¿Qué límite es ese? Ya lo hemos abordado antes. Observa que en la decisión que se ha realizado en el apartado anterior hemos partido de un intervalo de confianza del 95%. Ello implica un riesgo de error del 5%. Eso mismo es lo que podemos hacer aquí:

- Si la significación es menor que 0,05 entonces es asumible el riesgo y se rechaza la hipótesis nula.
- Si la significación es igual o mayor que 0,05 entonces el riesgo es elevado y se mantiene la hipótesis nula.

El umbral escogido recibe el nombre de “nivel de significación” o “riesgo α (alfa)”. α se decide en las mismas condiciones que abordamos para el nivel de confianza: valorando las consecuencias que se derivan de equivocarnos. Pero, también del mismo modo en que nos ocurrió entonces, ocurre en las PSHN que no es fácil valorar las consecuencias y que se arrastran hábitos. Generalmente el hábito establece un riesgo de 0,05 y, si se quiere ser más exigente, uno del 0,01.

Esquemas de decisión típicos

Existe un sinfín de situaciones en las que procede realizar una prueba de decisión. Aquí no podemos intentar la exhaustividad. Pero sí que podemos aspirar a cubrir muchas situaciones prácticas con un esquema sencillo. El objetivo es abordar situaciones en las que nos podemos encontrar y cómo abordarlas de manera correcta. Para ello, se muestran a continuación las situaciones en las que se han encontrado los grupos de trabajo de la clase.

Las situaciones en las que se han encontrado estos grupos se pueden clasificar en las que requieren abordar una relación entre dos variables que pueden ser nominales o cuantitativas. Frente a ello, vamos a escoger las pruebas de decisión no paramétricas, es decir, las que no exigen ningún comportamiento especial a los parámetros. Estas pruebas consideran variables ordinales más que cuantitativas. En tales casos nos situamos.

Dos variables nominales

Para abordar una posible relación entre dos variables nominales, primero hay que proceder a realizar una representación gráfica de diagrama de barras (por ejemplo), o construir una tabla de contingencia. Para este segundo caso, se procede a seguir la sucesión:

Analizar / Estadísticos descriptivos / Tablas de contingencia...

Una vez analizado en contenido de la tabla o de la representación gráfica, si muestra una posible relación, se procede a pedir un índice. En nuestro caso, recurrimos a la V de Cramer. Se consigue con el mismo procedimiento que la tabla de contingencia, pulsando sobre el botón Estadísticos.

La V de Cramer se mueve en el intervalo $(0,1)$, donde 0 representa ausencia total de relación, mientras que 1 significa relación máxima.

Por último, si la V de Cramer expresa una relación suficiente (no hay un patrón fijo, pero 0,4 comienza ya a tener algún significado, depende de otros aspectos, como la información que nos aporte visualizar la tabla de contingencia), entonces se procede a la inferencia. SPSS ya suministra automáticamente el grado de significación al pedir un índice de relación, como es el caso de la V de Cramer. Con el grado de significación ya sabemos qué hacer, gracias al apartado anterior.

Dos variables ordinales

Es el caso del ejemplo que hemos visto en el apartado anterior (relación entre las variables *¿Eres feliz?* y *¿Te sientes realizado/a?*) Como ya se ha mostrado, hay que comenzar con un diagrama de dispersión, para pasar a generar una rho de Spearman y, por último, un grado de significación. Del mismo modo que en el caso anterior, se pasa a la siguiente fase de análisis cuando la anterior así lo aconseja, es decir, cuando se observa que existe relación, aunque no sabemos cuánto o en qué medida es significativa.

Una variable ordinal y una nominal

El análisis previo comienza con un box-plot o diagrama de caja y patillas:

Analizar / Estadísticos descriptivos / Explorar...

En el cuadro de diálogo se activa únicamente "Gráficos". Se pulsa sobre el botón del mismo nombre y se accede a otro cuadro de diálogo donde se desactivan todas las opciones y se escoge "Niveles de los factores juntos".

Si el box-plot muestra relación, pasamos a la siguiente fase, que va a depender del número de categorías de la variable nominal: dos o más.

En el caso de dos niveles, se aplica la U de Mann-Whitney mediante:

Analizar / Pruebas no paramétricas / 2 muestras independientes...

Como se puede observar en el correspondiente cuadro de diálogo, ya se encuentra activada por defecto la U de Mann-Whitney, por lo que sólo hay que definir la variable de contraste (llamada habitualmente 'variable dependiente' en metodología) y la de agrupación (llamada habitualmente 'variable independiente'), definiendo el valor de cada uno de los dos grupos (codificación de la variable independiente).

En el caso de tres o más niveles de la nominal, se aplica la H de Kruskal-Wallis, con un procedimiento idéntico al anterior, pero escogiendo

Analizar / Pruebas no paramétricas / k muestras independientes...

Por lo demás, es análogo a la U de Mann-Whitney.

Ambos procedimientos generan índices que no sirven para reflexionar sobre la cuantía de la relación en la muestra, puesto que no tienen un valor máximo reconocible (varían de 0 a infinito). Sólo contamos con el box-plot previo y con la diferencia entre rangos. Como ya vimos en otro lugar, los rangos son órdenes o posiciones. Cuando las medias de rangos que aparece en los dos procedimientos son muy diferentes entre sí, entonces hay indicios de relación. Por lo demás, estas pruebas sirven únicamente para acceder al grado de significación.

Anexo: archivo de datos

El archivo de datos que se ha utilizado en los ejemplos de este documento parte de 6 variables y 84 casos. Los nombres de variables y sus etiquetas (con etiquetas de valores cuando procede) se muestran a continuación. Tras esta especificación se encuentran los datos.

feliz ¿Eres feliz? (respuesta de 0 a 10)
 realizado ¿Te sientes realizado/a? (respuesta de 0 a 10)
 infiel ¿Cuántas veces has sido infiel?
 destino ¿Hay un destino o el tiempo hace un camino?
 1: destino; 2: tiempo.
 nacido ¿Te arrepientes de haber nacido?
 1: sí; 0: no.
 objetivo ¿Cuál es el objetivo de tu vida?
 1: terminar la carrera.
 2: ser feliz.
 3: encontrar trabajo.
 4: tener dinero.
 5: ser libre.
 6: ser independiente.
 7: vivir dignamente.
 8: formar una familia.

feliz	realizado	infiel	destino	nacido	objetivo
4,0	5	0	2	0	
1,0	8	0	1	0	2
7,0	8	0	2	0	5
5,0	6	1		0	2
8,0	9	1	2	0	2
7,0	7	0		0	2

8,0	8	0		0	2
7,0	5	0	1	0	5
6,0	7	0	1	0	1
5,0	4	0	2	0	2
7,0	5	2	2	0	2
3,0	6	0	1	0	7
8,0	5	0	2	0	5
9,0	7	0	2	0	2
7,0	7	0	1	0	6
7,0	4	4	1	0	2
8,0	8	0	1	0	2
5,0	3	0	2	0	2
6,0	5	0	2	0	1
9,0	8	0	1	0	2
7,0	5	0		0	2
8,0	6	0		0	5
8,0	6	0	1	0	2
8,0	8	0	2	0	5
5,0	4	2	2	0	2
9,0	3	0	1	0	6
8,0	7	0	2	0	2
6,0	5	0	2	0	2
8,0	6	1	2	0	2
7,0	5	0	2	0	2
4,0	5	2		0	7
9,0	8	0	1	0	7
7,5	7	1	2	0	2
8,0	7	3	1	0	2
8,0	7	0	1	0	2
9,0	6	5	2	0	2
6,5	4	4		0	2
8,0	7	0	2	0	2
5,0	5	0	2	0	2
8,0	7	2	2	0	7
8,0	6	2	2	0	8
4,0	7	0	1	0	5
9,0	8	2	1	0	2
9,0	7	0	1	0	2
7,0	6	4		0	2
5,0	6	4	2	0	5
8,0	5	0	2	0	2
7,0	5	0	1	0	2
5,0	5	0	2	0	2
8,0	6	0	1	0	2
9,0	7	0	1	0	2
8,0	7	0	2	1	1
8,0	7	0	2	0	2
8,0	7	0	1	0	2
6,0	6	0	2	0	2
9,0	6	0	2	0	2
7,0	7	0		0	2
5,0	6	1	1	0	2
9,0	8	0	2	0	2
9,0	7	0	2	0	6
8,0	7	0	1		2
6,0	7	4	2	0	6
10,0	8	0	2	0	6
7,0	7	5	2	0	2
7,0	5	0	1	1	2
6,0	5	3	2	0	2
9,0	8	0	1	0	2
8,0	7	0	1	0	2
7,0	7	0	2	0	6
10,0	7	0	1	0	2

8,0	9	0	1	0	8
8,0	7	1	1	0	2
5,0	6	3	2	0	7
5,0	5	0	2	0	
8,0	6	0	2	0	7
10,0	6	0	1	0	2
	5	0	2	0	5
7,5	7	0	2	0	2
8,0	7	0	2	0	2
10,0	7	0	2	0	8
8,0	6	0	1	0	1
6,0	7	3	1	0	2
8,0	9	4	2	0	2
7,0	5	2	1	0	6