Para asegurar la exactitud metodológica y la relevancia y validez de los resultados, los grupos de investigación experimental en nuestro campo optan en los últimos tiempos por un acceso multimetodológico, una metodología que se denomina "triangulación". Recientemente ha sido publicado un volumen monográfico dedicado al tema: Alves, F. (ed.) (2003): *Triangulating Translation*. Amsterdam, John Benjamins Publishing Company.

Ejemplo: Pacte utiliza para la recogida de datos:

Traducciones de textos (traducción directa e inversa).

Translation protocols (utilizando los programas informáticos Proxy y Camtasia que permiten observar y grabar los comportamientos del sujeto durante el proceso de producción de la traducción).

Cuestionarios iniciales para seleccionar a los sujetos y como fuente de información al interpretar los resultados.

Cuestionarios para recopilar la opinión de los sujetos sobre lo que es una buena traducción.

TAPs retrospectivos con la grabación informática para recabar información complementaria.

Entrevistas.

Wordsmith tools para analizar el corpus de traducciones elaboradas por los traductores durante el experimento.

3.12. PROCEDER A LA RECOGIDA DE DATOS

Un buen consejo: limitémonos a recoger datos que sean relevantes para nuestro estudio, especialmente si se trata de trabajos académicos, como memorias, tesinas o tesis de doctorado. Aquellos que se dicen "recogeré también datos sobre tal y tal cosa, puede que me sirvan en algún momento" contravienen el principio de la *practicabilidad* o *economía científica*, aquel criterio descrito por Giegler (1974) que exige que los experimentos sean diseñados de manera lo más simple posible para evitar una sobrecarga de los probandos, para que sean manejables en su totalidad y para que el análisis de los resultados no implique un esfuerzo desmesurado por parte de los investigadores, como lo mencionamos más arriba.

Winston Churchill afirmaba: "Sólo me fío de las estadísticas que he manipulado yo mismo".

3.13. REALIZAR EL ANÁLISIS ESTADÍSTICO DE LOS DATOS

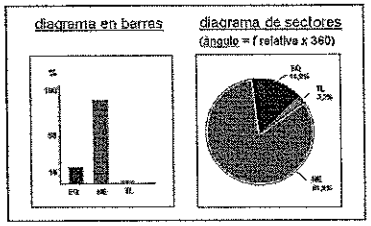
La estadística pone a nuestra disposición procedimientos y técnicas que permiten describir y analizar los datos obtenidos. Utilizamos los métodos de estadística descriptiva con la esperanza de que coincidan con las características básicas de la población (universo del que se ha retirado la muestra) y los métodos de estadística inferencial, que están basados en el cálculo de probabilidades y se ocupan de extender

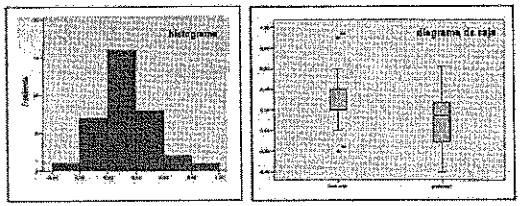
o extrapolar a toda una población las informaciones obtenidas de una muestra representativa. Aquí solo podemos presentar la 'filosofía' de algunas medidas y tests estadísticos que nos pueden ayudar en la descripción de nuestros datos; en todo caso, antes de emplear baterías de tests de estadística (como, por ejemplo, el SPSS, el *Statistical Package for Social Sciences*) se requiere la profundización en este campo o el apoyo de un especialista en metodología.

3.13.1 Métodos de estadística descriptiva

Cuando nuestros datos son categoriales (escala nominal) solemos describirlos mediante *porcentajes de frecuencia* (%) y representarlos gráficamente mediante tablas o diagramas. Hemos analizado, por ejemplo, los procedimientos que utilizan los alumnos al enfrentarse a metáforas culturales:

PROCEDIMIENTO	CASOS	PROPORCIÓN	PORCENTAJE
Equivalencia	60	0,149	14,9 %
Neutralización	331	0,819	81,9 %
Traducción literal	13	0,032	3,2 %
Total	404	1,0	100 %





Para hacernos una primera idea de la distribución de datos numéricos recurrimos normalmente al *histograma*, que es una representación gráfica de una variable en

Interesante:

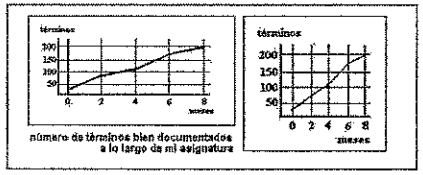
La aplicación de procedimientos estadísticos se remonta hacia el año 3050 a. C., cuando se efectuó en Egipto un registro de la población y la riqueza con el fin de preparar la construcción de las pirámides.



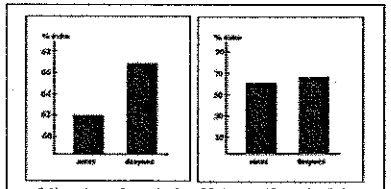
forma de barras donde la superficie de cada barra es proporcional a la frecuencia de los valores representados. El SPSS nos ofrece también un *diagrama de caja* que representa gráficamente la distribución de los colectivos y los valores atípicos.

Interesante: La representación gráfica da lugar a engaños (intencionados o no): por ejemplo, se pueden resaltar los resultados en la dirección deseada, atenuando o acentuando una tendencia:

mediante el corte en el eje horizontal



mediante el corte en el eje vertical



Medidas de tendencia central

Para la descripción de datos numéricos nos servimos de las medidas de tendencia central, posicionales o de dispersión, siempre esperando que representen la totalidad de la población de la que hemos retirado nuestra muestra.

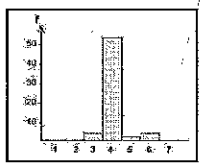
La medida central por excelencia es la *media*, que se calcula dividiendo los datos por el número de participantes.

$$\bar{x} = \frac{(x_1 + x_2 + \ldots + x_n)}{n} = \frac{\sum_{i=n}^{i=n} x_i}{n}$$

La *mediana* es el valor de X (abscisa) que corresponde a la recta vertical que divide un histograma en dos partes de igual área. *Cuartiles*, *deciles* y *percentiles* son análogos a la mediana, parten la muestra en cuatro, diez, cien partes

$$\tilde{X} = \frac{n+1}{2}$$

La *moda* es el valor más frecuente de la distribución, puede no existir, es decir, ningún valor de la distribución sobresale de manera evidente a los demás, e incluso no ser única en caso de existir.

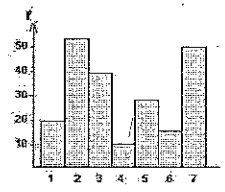


Interesante: Cuando media, mediana y moda no coinciden, nos encontramos ante una situación no esperada, no "normal" (en el ejemplo, ni media ni mediana ni moda reflejan la distribución real).

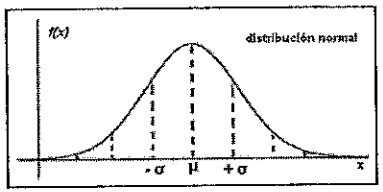
La media en esta distribución es: 4,02 (20 x 1 + 55 x 2 + 40 x 3 + 10 x 4 + 15 x 5 + 15 x 6 + 50 x 7 = 825 dividido por 205).

La mediana es 3 (205 + 1 dividido por 2 = 103, está en el valor 3).

La distribución es bimodal, tiene dos modas: 2 y 7.



Se puede hablar de normalidad en una distribución de datos cuando la media, la mediana y la moda coinciden en su valor, lo cual origina gráficamente una curva simétrica, donde su eje de simetría es el punto donde coinciden estas tres medidas de tendencia central. Esta *distribución normal* es conocida también como *campana o curva de Gauss*.



Medidas de dispersión

La *desviación estándar* (o típica) va representada por la letra griega sigma (σ) y es la raíz cuadrada de la varianza de una distribución. Es una medida de dispersión que representa el espacio de la curva normal comprendido entre la media y +/- σ.

$$\sigma = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}}$$

Si la desviación típica es pequeña, los datos están agrupados cerca de la media; si es grande, están muy dispersos; es una medida muy útil para tener una primera información sobre nuestros datos.

Medidas descriptivas: http://endrino.cnice.mecd.es/~jhem0027/estadistica/estadistica02.htm

Cálculo de tendencias de una muestra: http://www.elosiodelosantos.com/descriptiva.html

Ejemplo: distribución de la inteligencia (WAIS - Wechsler Adult Inteligence Scale)

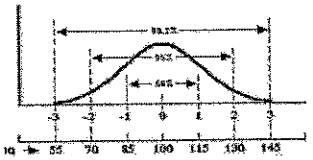
1 = 68 % de los casos = IQ entre 85 y 115

2 = 95 % de los casos = IQ entre 70 y 130

3 = 99,75 % de los casos = IQ entre 55 y 145

→ sólo el 0,13% de las personas tienen un IQ más allá de 145, pero no hay test que lo mida;

→ un cociente intelectual de 90 es, estadísticamente, tan "normal" como uno de 110.



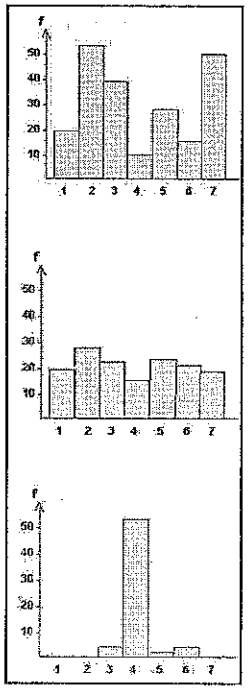
Inciso sobre la eliminación de ítems:

Es necesario eliminar ítems, por ejemplo, en el proceso de elaboración de un cuestionario, en los casos siguientes.

- Una muestra homogénea (por ejemplo traductores experimentados) arroja una distribución bimodal, (en contraposición a la distribución de Gauss, el histograma muestra dos modas), porque podemos inducir que el ítem está

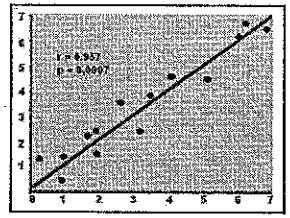
mal formulado o que el colectivo que responde al ítem carece de criterios para valorar.

- La desviación típica es muy alta (el histograma no muestra una moda clara), pues significa que las valoraciones de los sujetos están distribuidas a lo largo de casi toda la escala, es decir, la escala en cuestión no tiene el suficiente valor informativo o discriminatorio para nuestros fines.
- La desviación típica es muy baja, ya que esto significa que el ítem no permite discriminación (no sirve para nada), por ejemplo, si está mal formulado.
- Cuando los resultados de los diferentes grupos de los que suponemos tienen características claramente diferenciadas (por ejemplo, alumnos de primero y traductores expertos) son iguales, podemos inducir que el ítem en cuestión mide obviedades (siempre y cuando no nos interese ver justamente en qué coinciden todos los encuestados).



La correlación

El *coeficiente de correlación de Pearson (r)* es un índice que mide la relación lineal entre 2 variables cuantitativas. Indica en qué grado 2 variables X e Y fluctúan simultáneamente, es decir, cuánto aumenta X al aumentar Y *(correlación positiva)*, o cuánto aumenta X al disminuir Y *(correlación negativa)*. Exige una distribución normal de las variables. A partir de unas tablas se obtiene el valor de significación (p), que depende del grado de correlación entre ambas variables y del tamaño de la muestra. Para cumplir el postulado de "distribución normal" tenemos, por ejemplo, la "transformación Z" de Fisher, que transforma los datos en bruto (de una escala de intervalos) en una distribución con la media <x> = 0 y la desviación típica σ = 1, es decir, en una curva de Gauss.



Cálculo de regresión (*i.e.* Pearson r): http://www.elosiodelosantos.com/regresionlineal.html

El *coeficiente de correlación de rangos de Spearman (rho)* es una medida de asociación lineal que utiliza los rangos, números de orden, de cada grupo de sujetos y compara

dichos rangos. El rho de Spearman, llamado también r de Spearman, es exactamente lo mismo que el coeficiente de correlación de Pearson calculado sobre el rango de observaciones. En definitiva la correlación estimada entre X e Y se halla calculando el coeficiente de correlación de Pearson para el conjunto de rangos de dos grupos apareados. Por ejemplo, si queremos saber si los buenos traductores en traducción directa también serán buenos en inversa, ordenamos a los sujetos por los resultados de traducción directa e inversa y calculamos el rho de Spearman. Si el coefficiente es alto vemos confirmada nuestra idea.

El valor del coeficiente de correlación se altera de forma importante ante la presencia de un valor extremo, como sucede con la desviación típica. Ante estas situaciones conviene realizar una transformación de datos que cambia la escala de medición y modera el efecto de valores extremos (como la transformación logarítmica), o bien eliminar estos sujetos de la muestra si sospechamos por algún motivo que no pertenecen al universo del que hemos extraído nuestra muestra.

¡Ojo!

La correlación elevada y estadísticamente significativa no tiene que asociarse a causalidad. La causalidad es un juicio de valor que requiere más información que un simple valor cuantitativo de un coeficiente de correlación.

Cuando confirmamos objetivamente que dos variables están correlacionadas pueden haber diferentes motivos para esta correlación: a) puede ser que X influencie o cause Y, b) puede que Y influencie o cause X, c) X e Y pueden estar influenciadas por terceras variables que hace que se modifiquen ambas a la vez. Las mencionadas cigüeñas de Hamburgo son un buen ejemplo de esta falacia, denominada "correlación espuria", que Wikipedia define de la siguiente manera: "En estadística una correlación espuria es una relación matemática en la cual dos acontecimientos no tienen conexión lógica, aunque se puede implicar que la tienen debido a un tercer factor no considerado aún (llamado "factor de confusión" o "variable escondida"). La relación espuria da la impresión de la existencia de un vínculo apreciable entre dos grupos que es inválido cuando se examina objetivamente."

3.13.2. Métodos de estadística inferencial

La estadística inferencial se ocupa de extrapolar a toda una población, partiendo de cálculos probabilísticos, informaciones obtenidas de una muestra representativa. Existen muchos tests para saber si los datos obtenidos tienen alguna significancia a nivel estadístico. El programa SPSS (*Statistical Package for Social Sciences*) nos es aquí de gran ayuda. En ciencias sociales se suele fijar el nivel de significancia en 5%, es decir, las tendencias observadas solo se pueden atribuir en el 5% de los casos al azar. Claro que esto implica la posibilidad de cometer errores: cometemos, por ejemplo, un error de tipo 1 (error α) cuando aceptamos que existe una diferencia entre nuestros grupos experimentales (rechazo de la hipótesis nula) cuando en realidad los dos grupos son iguales, o bien, un error de tipo 2 (error β) cuando, dado que nuestros datos no arrojan una significancia, consideramos que los dos grupos son iguales (aceptamos la hipótesis nula) cuando en realidad son diferentes.

Los tests de estadística inferencial se dividen en *tests no paramétricos*, que requieren

datos numéricos solo de nivel escala ordinal, y *tests paramétricos,* que requieren distribución normal y escala de intervalos. A continuación, un poco de 'filosofía' de algunos tests seleccionados:

La prueba del *Ji-cuadrado de Pearson* (χ^2 o también Chi-cuadrado) se utiliza para saber si los datos de una muestra son conformes a una distribución teórica que sospechamos que es la correcta (por ejemplo, la distribución normal requerida para aplicar tests paramétricos), o también para contrastar la igualdad de procedencia de un conjunto de muestras de tipo cualitativo, por ejemplo, para saber si dos muestras extraídas de un mismo universo son realmente 'iguales' o 'paralelas'.

La idea de la que parten los tests no paramétricos *U de Mann-Withney* y *Suma de Rangos de Wilcoxon* es la siguiente: si unimos las dos muestras (muestra X, valores x_1 ..... x_i, muestra Y, valores y_1 ..... y_i) y ponemos los valores en orden, la alternancia entre las x_i, y las y_i, debería ser bastante regular. Si no se da esta regularidad, podemos inferir que las dos muestras difieren, es decir, que la influencia de algo (de alguna variable) provoca estos resultados.

La *prueba t de Student* compara las medias de dos categorías dentro de una variable dependiente, o las medias de dos grupos diferentes. Es una prueba paramétrica, o sea, que solo sirve para comparar variables numéricas de distribución normal. Se utiliza especialmente para comparar las medias de un mismo grupo en diferentes etapas (*repeated measurement*).

La *prueba W de Kendall* (coeficiente de concordancia) nos indica, por ejemplo, si las evaluaciones de traducciones por parte de diferentes profesores concuerdan o no.

El *α de Cronbach* nos informa sobre la consistencia interna, la fiabilidad (*reliability*) de una escala. El α debe tener, por lo menos, un valor de 0,8 para ser considerado fiable.

El *Análisis de varianza* es una prueba semejante a la prueba t de Student en cuanto a la práctica y, básicamente, se utiliza para corroborar si la significación de diferencias entre medias de dos o mas grupos en mediciones sucesivas son o no debidas al azar. Es indicado por ello en estudios de *repeated measurement.*

El *Análisis factorial* se utiliza para el estudio e interpretación de las correlaciones entre un grupo de variables. Parte de la idea de que dichas correlaciones no son aleatorias, sino que se deben a la existencia de factores comunes entre ellas. El objetivo del análisis factorial es la identificación y cuantificación de dichos factores comunes.

¡Ojo!

El término "estadísticamente significativo" se percibe cada vez en mayor medida como un sello de "garantía de calidad", pero no tiene en consideración la relevancia de un estudio, que va más allá de cálculos numéricos.

3.14. Interpretar los resultados

El último paso en un trabajo científico es contrastar los resultados con las hipótesis para corroborar nuestra idea o rechazarla, lo que nos obligaría a modificar nuestra hipótesis o teoría, o nuestro modelo, una modificación que se volverá a validar en un proceso empírico, con lo que se cierra la "rueda de la ciencia".

3.15. Comunicación de los resultados a la comunidad científica

La forma usual de comunicar los resultados de la investigación a la comunidad científica es de manera oral mediante ponencias o comunicaciones en congresos especializados o mediante informes o artículos publicados en revistas o volúmenes monográficos dedicados a nuestro campo científico.

Páginas web que recogen los congresos de traducción: http://www.iatis.org/content/iatis2006/index.php

http://www.tolk.su.se/konfindx.html

Página web que recoge la evaluación CARHUS de revistas de traducción: http://www10.gencat.net/dursi/ca/re/aval_rec_sist_pc05_filologia_traduccio.htm