

En **Fisioterapia** trabajamos con datos de variables que pueden ser **nominales, ordinales, de intervalo o de razón** (<http://www.fisioeducacion.net/fisios/fisioinvestigacion/206-medicion-y-escalas-de-medida>). Estos datos pueden provenir de la medición directa de estas variables (escala visual analógica, fuerza muscular, distancia dedos suelo, amplitud articular, etc.) o de frecuencias que provienen de un proceso de conteo (número de pacientes tratados, número de técnicas realizadas, número de fisioterapeutas en un curso, etc.). Una vez que el investigador recaba la información mediante la recogida de datos, dispone de un listado de los mismos. Si hay pocos valores, un simple vistazo nos serviría para describir el fenómeno estudiado, pero esto no es lo más habitual. Lo normal, es que dispongamos de una cantidad de datos que necesiten ser organizados mediante una distribución de frecuencias.

La distribución de frecuencias es la agrupación de datos en categorías mutuamente excluyentes que indican el número de observaciones de cada categoría. La distribución de frecuencias cumple con varias funciones: organiza los datos, ofrece la información necesaria para realizar representaciones gráficas y facilita los cálculos para obtener los estadísticos muestrales.

Por ejemplo, si queremos hacernos una idea precisa de los hombres y mujeres que intervienen en un estudio sobre incontinencia urinaria de un grupo de 84 personas, necesitamos construir una distribución de frecuencias.

Variable X	n_i	p_i	P_i
Hombre	16	0,2	20
Mujer	68	0,8	80
	n=84	1	100

Ahora conocemos el número de observaciones en cada categoría, **la frecuencia absoluta** (n_i). También conocemos **la frecuencia relativa o proporción de cada categoría** (p_i), que se obtiene dividiendo la frecuencia absoluta entre el número total de de observaciones (n). La frecuencia relativa también se expresa en términos de **porcentaje** (P_i) para lo cual multiplicamos cada una de las proporciones por cien.

En el anterior ejemplo, hemos trabajado con una variable nominal. Si trabajásemos con una variable ordinal, podríamos añadir a la tabla tres columnas más: **la frecuencia absoluta acumulada** (n_a), **la frecuencia relativa acumulada o proporción acumulada** (p_a) y **el porcentaje acumulado** (P_a). La frecuencia absoluta acumulada es el número de veces que se repite cada modalidad y cualquiera de las modalidades inferiores. la proporción acumulada o frecuencia relativa acumulada es el cociente entre la frecuencia absoluta acumulada de cada clase y el

Investigación

Investigación
Jueves, 02 de enero de 2014
Gema Gallardo

total de observaciones. El porcentaje acumulado es el valor de la frecuencia relativa acumulada multiplicado por cien. Por ejemplo, utilizando la variable *nivel de entrenamiento*:

X	n_i	p_i	P_i	n_a	p_a	P_a
Nulo	13	0,325	32,5	13	0,325	32,5
Bajo	15	0,375	37,5	28	0,7	70
Medio	8	0,2	20	36	0,9	90
Alto	4	0,1	10	40	1	100
	40	1	100			

Cuando trabajamos con variables cuantitativas puede ocurrir que el número de valores que tome la variable sea muy amplio. Por ejemplo, con la variable edad, en un estudio sobre el perfil del paciente que acude a una unidad de fisioterapia hospitalaria tendríamos una lista muy extensa, por lo que nos sería muy útil la agrupación en **intervalos**, que consiste en formar grupos de valores consecutivos de la variable. Elegiríamos cuantos intervalos vamos a utilizar, teniendo en cuenta el valor más bajo y el valor más alto de nuestra variable, es decir, el paciente más joven y el paciente más viejo que hemos tratado en nuestra unidad. Podemos optar por hacer muchos intervalos de poca **amplitud** (diferencia entre los límites del intervalo) o bien, pocos intervalos de mucha amplitud, intentando equilibrar la precisión de los datos con la manejabilidad de los mismos. Quizás no sería muy útil utilizar intervalos de poca amplitud en unidades donde todos los pacientes rondan una edad, pero sí, en unidades donde se tratan niños, adolescentes, adultos y tercera edad. También podemos utilizar **intervalos abiertos**, " 80 o más".

Según esta tabla de distribución de frecuencias agrupadas la variable edad no puede tomar valores entre 40 y 41 , lo cual no es cierto en la práctica. Debemos considerar estos límites como informados o aparentes. Los límites exactos del intervalo serían "20,5-40,5" y el siguiente "40,5-60,5", de manera que el límite superior exacto de un intervalo coincida con el límite inferior exacto del siguiente.

Intervalos X			n_i	p_i	n_a	p_a
20	10	0,2	10	0,2		
años		50		50		
o						
men						
os						
21-	3	0,0	13	0,3		

Investigación

Investigación
Jueves, 02 de enero de 2014
Gema Gallardo

40 años		75		25
41-60 años	13	0,325	26	0,650
61-80 años	7	0,175	33	0,825
81 o más	7	0,175	40	1
	40	1		

En un próximo artículo os contaremos cómo se representa gráficamente una distribución de frecuencias.