

12 Anexo; Definición de la muestra y obtención de la muestra

Para iniciar el estudio en primer lugar se determina cual es la población que se va a analizar. Las características escogidas para la elección y análisis de los datos son: en primer lugar la titulación ya que se considera una variable muy importante porque permite que se puedan sacar conclusiones para las diversas tipologías de titulaciones que hay en estos momentos en la UPV. En segundo lugar, el curso de finalización (desde 95/96 hasta 99/00) se considera una característica relevante por el hecho que puede permitir ver la evolución a lo largo del tiempo de las diversas actuaciones que ha realizado la Universidad en este período, así como cambios en los egresados. Y por último el género, ya que se pueden estudiar si existen diferencias entre los hombres y las mujeres titulados en esta Universidad. Estos datos junto con otras características de los individuos como población, fecha de nacimiento, etc... se obtuvieron por medio del Centro de Procesos de Datos de esta Universidad.

Una vez decidido las características a estudiar se pasó a realizar una depuración de los datos. Esta consiste en eliminar de la población a todos aquellos individuos que no se les va a poder localizar porque no disponen de teléfonos de contacto. Una vez hecho esto se contó con una población inicial de 16991 titulados, distribuidos según la tabla adjunta:

TITULACIÓN	95/96	96/97	97/98	98/99	99/00	TOTAL
ARQUITECTURA	79	100	136	245	247	807
ARQUITECTURA TÉCNICA	178	270	311	240	258	1257
DIPLOMATURA DE INFORMÁTICA +1	79	75	51	305	23	533
INGENIERO T. EN INFORMÁTICA +1	56	173	194	243	202	868
ING. CAMINOS CANALES Y PUERTOS	34	26	47	82	88	277
ING. EN AUTOMÁTICA Y ELEC. INDUST.	0	1	12	23	20	56
ING. EN GEODESIA Y CARTOGRAFÍA	0	34	32	30	19	115
ING. DE ORGANIZACIÓN INDUST. (VAL.)	0	22	48	50	63	183
ING. DE ORGANIZACIÓN INDUST. (ALC.)	1	15	24	22	28	90
INGENIERIA INDUSTRIAL	179	222	364	368	332	1465
INGENIERIA QUÍMICA	0	0	14	27	75	116
ING. TÉCNICA AGRÍCOLA	276	284	266	267	271	1364
ING. TÉC. DISEÑO INDUSTRIAL (VAL.)	0	9	21	53	68	151
ING. TÉC. DISEÑO INDUSTRIAL (ALC.*)	0	0	0	14	23	37
ING. TÉCNICA TOPOGRÁFICA	34	74	65	83	110	366
ING. TÉCNICA FORESTAL	1	17	29	53	64	164
ING. TÉCNICA INDUSTRIAL (VAL.)	481	522	680	442	501	2626
ING. TÉCNICA INDUSTRIAL (ALC.)	132	159	210	90	130	721
ING. TÉC. TELECOMUNICACIÓN (GAN)+2	0	14	25	55	89	183
ING. TÉC. TELECOMUNICACIÓN (ALC)+2	4	20	46	48	62	180
ING. TÉCNICA OBRAS PÚBLICAS	40	63	67	103	121	394
INGENIERÍA AGRONÓMICA	80	101	132	121	203	637
INGENIERÍA TELECOMUNICACIONES	160	175	196	206	212	949
INGENIERÍA INFORMÁTICA +3	0	0	46	129	104	279
LDO. INFORMÁTICA +3	283	245	317	315	556	1716
LDO. BELLAS ARTES	264	344	295	251	245	1399
LDO. CC. AMBIENTALES (VAL.)*	0	0	0	2	21	23
LDO. CC. AMBIENTALES (GAN.)*	0	0	0	0	13	13
LDO. EN DOCUMENTACIÓN*	0	0	0	10	12	22
TOTAL	2361	2965	3628	3877	4160	16991

Las titulaciones señaladas (*) se ha decidido no incluirlas en este estudio, a la espera de que el número de titulados aumente y poder estudiar su verdadera influencia en el entorno socioeconómico ya que son titulaciones de creciente creación y el número de alumnos titulados no está estabilizado en el mercado laboral, siendo estos datos poco significativos de la realidad. Las titulaciones que están señaladas con (+) y un número son estudiadas conjuntamente, puesto que son carreras que se pueden considerar equivalentes, esto permite tener un mayor número de individuos por grupo.

Una vez conocida la población objeto de estudio se plantea el tipo de muestreo que se va a utilizar. Este va estar condicionado según las variables de selección determinadas con anterioridad. Las variables año de titulación, titulación y género, permiten la creación de grupos homogéneos de individuos donde es previsible obtener respuestas similares entre ellos y diversas entre los distintos grupos. Por ello se plantea el trabajo con un muestreo aleatorio por estratos, donde cada uno de los estratos o subconjuntos los proporcionan las variables referidas con anterioridad.

El sistema utilizado para la realización de las entrevistas ha sido el CATI (entrevista telefónica asistida por ordenador), permite seleccionar de forma aleatoria la muestra así como aplazar entrevistas para alcanzar la máxima bondad de la muestra ya que no son rechazados los individuos de máxima dificultad de localización y lleva de forma automática el control de las cuotas que establecen los estratos, justificando la utilización de un muestreo aleatorio simple dentro de cada estrato como el óptimo.

La selección del tamaño de los estratos se hace por afijación proporcional al tamaño de cada titulación. Para estudiar la evolución se hace una afijación según el número de titulados en cada año. Y finalmente se estratifica, de forma proporcional al género, con el objetivo de poder sacar conclusiones válidas, no sólo de carácter general para la totalidad de la población objeto de este estudio, sino también de carácter más concreto por titulación, género y año de finalización.

Para elegir el tamaño de la muestra general se plantea su cálculo a partir de un error de muestreo fijado que proporcione unas buenas estimaciones, y después se calcula el tamaño de la muestra. Para que las estimaciones sean precisas se considera un error fijo del 2% para el cual obtenemos un tamaño muestral de aproximadamente 2000 entrevistas. Como se comenta con anterioridad se selecciona el número de individuos para cada uno de los estratos de manera proporcional al tamaño de este en la población, técnica denominada de afijación proporcional. La expresión que se utiliza en este caso para el cálculo es:

$$n_h = N_h \frac{n}{N} \quad (1)$$

N= Tamaño de la población.

n= Tamaño de la muestra.

N_h = Tamaño poblacional del estrato h-ésimo.

n_h = Tamaño muestral del estrato h-ésimo.

L= Número de estratos.

$$\text{TOTAL} \quad n = \sum_{h=1}^L n_h = 2000$$

Como se ha indicado antes, una vez fijado el error global y calculado el tamaño muestral se calculan los tamaños muestrales para los estratos y sus errores.

La expresión utilizada para conocer el tamaño muestral total, sustituyendo n_h según (1), es:

$$e^2 = k^2 \frac{1}{N^2} \sum \frac{N_h^2 (N_h - n_h)}{N_h - 1} \frac{P_h Q_h}{n_h} = k^2 \left(\sum W_h^2 \frac{N_h}{N_h - 1} \frac{P_h Q_h}{n_h} - \frac{1}{N^2} \sum \frac{N_h^2}{N_h - 1} P_h Q_h \right) \quad (2)$$

donde:

k = Coeficiente de confianza.

N = Tamaño de la población.

N_h = Tamaño poblacional del estrato h -ésimo.

n_h = Tamaño muestral del estrato h -ésimo.

P_h = Estimador de la proporción del estrato h -ésimo.

$Q_h = 1 - P_h$.

$W_h = N_h / N$.

Al utilizar el muestreo aleatorio simple, para el cálculo de los errores en los estratos individualmente se utiliza:

$$e_h^2 = k^2 \frac{N_h - n_h}{N_h} \frac{P_h Q_h}{n_h}$$

Una vez se han realizado las operaciones pertinentes según (2) y tomando una estimación máxima de la proporción para cada uno de los estratos, se obtiene un tamaño de muestra total de 2051 individuos.

Hay que tener en cuenta que los errores se han proporcionado con cotas máximas del error para el nivel de confianza previsto (95 %), ya que para estimar las proporciones se ha utilizado su valor máximo. Si se tiene en cuenta que en algunas preguntas, como si esta trabajando en estos momentos, la proporción de individuos que responden afirmativamente es más de 85%, se puede afirmar que los errores son inferiores a los calculados. La muestra final queda reflejada en la siguiente tabla.

TITULACION ☐	TOTAL ☐
ARQUITECTURA ☐	98 ☐
ARQUITECTURA TECNICA ☐	154 ☐
DIP. E ING. TÉC. DE INFORMÁTICA ☐	159 ☐
ING. CAMINOS, CANALES Y PUERTOS ☐	34 ☐
ING. EN AUTOMATICA Y ELEC. INDUST. ☐	7 ☐
ING. EN GEODESIA. Y CARTOGRAFIA. ☐	14 ☐
ING. DE ORGANIZACIÓN INDUS.(VAL) ☐	21 ☐
ING. DE ORGANIZACIÓN INDUS.(ALC) ☐	11 ☐
INGENIERIA INDUSTRIAL ☐	182 ☐
INGENIERIA QUÍMICA ☐	13 ☐
ING. TÉCNICA AGRÍCOLA ☐	164 ☐
ING. TÉC. DISEÑO INDUSTRIAL(VAL) ☐	18 ☐
ING. TÉCNICA TOPOGRAFICA ☐	43 ☐
ING. TÉCNICA FORESTAL ☐	34 ☐
ING. TÉCNICA INDUSTRIAL (VAL) ☐	307 ☐
ING. TÉCNICA INDUSTRIAL (ALC) ☐	84 ☐
ING. TÉC. TELECOMUNICACION ☐	44 ☐
ING. TÉCNICA OBRAS PUBLICAS ☐	45 ☐
INGENIERIA AGRÓNOMICA ☐	79 ☐
INGENIERIA TELECOMUNICACIONES ☐	116 ☐
LDO. Y INGENIERIA INFORMATICA ☐	247 ☐
LDO. BELLAS ARTES ☐	177 ☐
TOTAL ☐	2051 ☐