



## RELACIÓN DE PROBLEMAS nº 2

### ESTADÍSTICA DESCRIPTIVA BIIDIMENSIONAL

1. Sean las siguientes variables:

X: Hábito de fumar (SI para el fumador y No para el no fumador)

Y: Años de vida una vez detectado un tipo de cáncer.

Z: Exposición a rayos ultravioleta (que puede ser ALTA o BAJA)

¿Qué coeficiente deberá utilizarse para medir la relación entre cada dos variables de las anteriores?  
Razona la respuesta.

2. Algunas personas piensan que se trabaja mejor por la mañana, mientras que otros piensan que es por la tarde cuando se trabaja mejor. Se ha considerado el turno de 20 trabajadores en la mañana (1) y tarde (0) y además se han obtenido datos de la calidad del trabajo producido por ellos, basados en una escala de 100 puntos. Los datos son los siguientes:

Turno	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Calidad	65	80	55	60	55	70	60	70	55	70	40	70	50	40	60	50	40	50	40	60

Determinar, razonadamente, la relación que pudiera haber entre el turno y la calidad del trabajo producido.

3. La siguiente tabla muestra los criterios de preferencia seguidos por el psicólogo industrial y el jefe del departamento comercial de un grupo editorial referidos a 10 aspirantes a ocupar diversos puestos de trabajo:

Psicólogo	A	B	C	D	E	F	G	H	I	J
Jefe	D	A	B	F	G	C	J	I	E	H

Elegir, calcular e interpretar el coeficiente más adecuado para medir el acuerdo, o desacuerdo, entre los dos criterios de selección.

4. Se desea saber si existe relación entre dos asignaturas, X e Y en cuanto a la fecha de presentación al examen, Junio o Septiembre (considerando que no pueden presentarse en las dos convocatorias). Para ello se analiza una muestra de 600 alumnos, encontrándose que de todos ellos un 20% se presentan a ambas asignaturas en Junio, y un 10% se presentan a ambas asignaturas en Septiembre. Si en la signatura Y se han presentado el doble de alumnos en Septiembre que en Junio, elegir, calcular e interpretar el coeficiente más adecuado para estudiar la relación.

5. La desviación típica de un determinado grupo de personas en la variable ansiedad (X) es 2. Sabemos que la media en esta variable en el grupo de hombres es 10 y en el grupo de mujeres es de 5, sabiendo que el índice de asociación entre las variables de ansiedad y sexo es igual a (+1), y que el número de hombres es superior al de mujeres

- ¿Qué coeficiente habrá sido utilizado?
- Interpretar el valor de ese coeficiente
- Calcular la proporción de hombres que componen nuestra muestra.

6. En un grupo de estudiantes de Estadística, un test de razonamiento numérico X, con media 24 y desviación típica 14, explica el 64% de la varianza de las calificaciones en el examen final de la asignatura, Y, con media de 18 puntos y varianza 49. Calcular:

- La correlación entre las dos variables, sabiendo que  $r_{xy} > 0$
- La calificación que pronosticaremos a un alumno que en el test obtiene 10 puntos.
- La puntuación en el test de un alumno que en el examen está a una desviación típica de la media.

7. Con los datos de la siguiente tabla, en la que se recogen las puntuaciones de 7 sujetos en dos variables X e Y, calcular:

- El coeficiente de correlación de Pearson.
- La recta de regresión de Y sobre X en puntuaciones directas.
- La varianza de Y, la varianza de las puntuaciones pronosticadas y la varianza de los errores.

X	6	3	7	5	4	2	1
Y	7	6	2	6	5	7	2

8. Un descuidado investigador está consultando un listado incompleto de los resultados de un experimento realizado hace tiempo. Deduce que la ecuación de regresión es  $Y' = 3 + 0,2 X$  y que la media de X es 20. También tiene una tabla en la que anotó la media y la varianza de las siguientes 4 variables: X, Y, Y', Y-Y' pero además de estar incompleta utilizó una notación distinta y no recuerda el nombre que dio a cada una de esas variables. La tabla es:

	VD	P	E	VI
Media			0	20
Varianza	4	2,56	1,44	

Contestar las siguientes cuestiones:

- Puntuación pronosticada en Y para  $X = 22$  en puntuaciones directas y en puntuaciones de desviación.
- Identificar, razonadamente, cada una de las variables que aparecen en la tabla.
- Averiguar el valor de los datos que faltan en la tabla.

9. Se sabe que para los niños entre 1 y 10 años existe una relación lineal entre el peso (en kg) y la altura (en cms.), de tal manera que a un niño con un peso de 10 kg le corresponde una altura de 50cms. y a un niño de 30kg le corresponde una altura de 120cms. Con estos datos:

- ¿Cuál es la ecuación de regresión, en puntuaciones directas y de desviación, que permite pronosticar la altura (Y) a partir del peso (X) de un niño?
- Conociendo  $S_y = 5,25$  y  $S_x = 1,2$  ¿Cuánto vale el coeficiente de correlación de Pearson entre la altura y el peso?
- ¿Cuánto vale la desviación típica de los errores en los pronósticos en la altura a partir del peso?

10. En la siguiente tabla aparecen los datos de un experimento de estimación de pesos. En este tipo de tareas el sujeto debe sopesar los objetos que se le presentan y proporcionar una estimación del peso de los mismos. La variable X representa el peso real (en kg), mientras que la variable Y representa los valores numéricos proporcionados por un observador.

X	1	3	4	6	8
Y	1	2	4	4	5

- Ajustar una recta de regresión, por el método de mínimos cuadrados, a los datos de la tabla utilizando X como variable dependiente.
- Determinar la covarianza de X e Y utilizando para ello  $r_{xy}$
- Estimar, según la recta de regresión, el peso real que tendrá un objeto para el que se haya proporcionado una estimación de 2 unidades y su error de estimación.