

EXAMEN DE RECUPERACION BIOESTADISTICA II I TERMINO 2015

1. La siguiente tabla presenta cuatro conjuntos de datos para cuatro modelos de regresión lineal distintos. Las observaciones bajo el encabezamiento x(a)(b)(c) son los valores de una variable explicativa que es común en las tres regresiones con variables respuesta y(a), y(b) e y(c). Por último, y(d) y x(d) definen conjuntamente otra regresión.

Observación	x(a)(b)(c)	y(a)	y(b)	y(c)	x(d)	y(d)
1	10	8.04	9.14	7.46	8	6.58
2	8	6.95	8.14	6.77	8	5.76
3	13	7.58	8.74	12.74	8	7.71
4	9	8.81	8.77	7.11	8	8.84
5	11	8.33	9.26	7.81	8	8.47
6	14	9.96	8.10	8.84	8	7.04
7	6	7.24	6.13	6.08	8	5.25
8	4	4.26	3.10	5.39	19	12.50
9	12	10.84	9.13	8.15	8	5.56
10	7	4.82	7.26	6.42	8	7.91
11	5	5.68	4.74	5.73	8	6.89

Para cada uno de los cuatro casos mencionados:

- Obtener la recta de regresión de Y sobre X.
- Obtener el coeficiente de correlación y de determinación entre los pares de variables correspondientes.
- Dibujar la nube de puntos y comentar los resultados anteriores.

2. Se realiza un estudio de fotoperíodo en aves acuáticas. Se pretende establecer una ecuación mediante la cual pueda predecirse el tiempo de reproducción, Y, en base al conocimiento del fotoperíodo (número de horas de luz por día) bajo el que se inició la reproducción, X. Se obtuvieron datos del comportamiento de 11 *Aythya* (patos buceadores). Los resultados fueron los siguientes:

Tiempo reproducción	110	54	98	50	67	58	52	50	43	15	28
Fotoperíodo	12.8	13.9	14.1	14.7	15.0	15.1	16.0	16.5	16.6	17.2	17.9

- Dibujar la nube de puntos de Y = Tiempo de Reproducción sobre X = Fotoperíodo. A la vista de la misma contestar razonadamente a las siguientes preguntas: ¿Es razonable considerar la existencia de una relación lineal entre las variables? ¿Cuál sería el valor esperado del coeficiente de correlación, cercano a 1, a -1, o, a 0?
- Utilizando los gráficos de residuos tipificados que consideres más adecuados comenta el cumplimiento de las hipótesis de normalidad y homocedasticidad (constancia de las varianzas).
- ¿Cuál es la recta de regresión lineal que permite predecir el tiempo de reproducción a partir del fotoperíodo?
- Suponiendo que el modelo lineal es aplicable para describir la relación entre el tiempo de reproducción y el fotoperíodo, ¿podemos afirmar que las dos variables están efectivamente correlacionadas? Utilizando software estadístico formula y resuelve un contraste de hipótesis que te permita contestar a esta pregunta.

- e) Da estimaciones puntuales de los siguientes parámetros poblacionales β_0 , β_1 , ρ y $\sigma_{Y|X}$
- f) A partir del coeficiente de determinación y de su relación con la desviación típica residual valora la idoneidad del modelo lineal obtenido para describir la relación entre el tiempo de reproducción y el fotoperíodo. ¿Qué porcentaje de la variabilidad observada en los tiempos de reproducción no quedan justificados por las diferencias observadas en el fotoperíodo? ¿Qué porcentaje de las diferencias observadas en los tiempos de reproducción quedan explicados por la recta de regresión obtenida i.e., por las diferencias observadas en el fotoperíodo?
- g) ¿Qué cambio cabe esperar en el tiempo de reproducción de las aves acuáticas por cada hora extra de luz diaria en el momento de iniciarse la reproducción?
- h) ¿Cuál sería el tiempo de reproducción de un ave cuyo fotoperíodo ha sido de 14.5 horas? Da un intervalo de confianza al 95% para esta predicción.
- i) Considerando la población condicional formada por los tiempos de reproducción de las aves cuyo fotoperíodo es de 14.5 horas, da, para estas aves, estimaciones puntuales del tiempo medio de reproducción y de la varianza. Obtén un intervalo de confianza al 95% para dicha media.
- j) Compara los intervalos de confianza obtenidos en los dos apartados anteriores y explica la relación que crees hay entre ellos.
- k) ¿Es posible utilizar la recta de regresión obtenida para realizar una estimación del tiempo de reproducción de un ave cuyo fotoperíodo es de 20 horas? Justifica la respuesta.

3. La siguiente tabla muestra dos conjuntos de datos (X, Y_1) y (X, Y_2) . Los valores de X son los mismos para cada conjunto de datos, pero los valores de Y son diferentes.

X	Y_1	Y_2	X	Y_1	Y_2
0.61	0.88	0.96	2.56	1.97	1.20
0.93	1.02	0.97	2.74	2.02	3.59
1.02	1.12	0.07	3.04	2.26	3.09
1.27	1.10	2.54	3.13	2.27	1.55
1.47	1.44	1.41	3.45	2.43	0.71
1.71	1.45	0.84	3.48	2.57	3.05
1.91	1.41	0.32	3.79	2.53	2.54
2.00	1.59	1.46	3.96	2.73	3.33
2.27	1.58	2.29	4.12	2.92	2.38
2.33	1.66	2.51	4.21	2.96	3.08

- a) Calcular el coeficiente de correlación para cada conjunto de datos.
- b) Para el conjunto de datos (X, Y_1) , multiplicar los valores de X por 10 y multiplicar los valores de Y_1 por 3 y añádeles 5. Recalcular el coeficiente de correlación y compararlo con el valor obtenido antes de la transformación. ¿Cómo afecta una transformación lineal al coeficiente de correlación?
- c) Comprobar que las rectas de regresión para ambos conjuntos de datos son iguales.
- d) Dibujar las correspondientes nubes de puntos.

4. Para estudiar el efecto de la *iluminación* (A=natural, B=muy fuerte, C=escasa) en la *REPRODUCCION de un pez* se realiza un experimento que consiste en contar el número de ovocitos depositados, con un fondo de acuario de diferente color (b=blanco, c=cafe, v=verde) y diferente tamaño de acuario (g= grande, p= pequeño, n=normal). Obteniéndose los resultados expuestos en la siguiente tabla:

Tipo de papel			
ACUARIO	cafe	blanco	verde
grande	258 A	230 C	240 B
normal	235 B	270 A	240 C
pequeño	220 C	225 B	260 A

Analizar estos datos y estudiar la posible influencia de los factores *iluminación, color de fondo y diferente tamaño de acuario* en la variable de interés *REPRODUCCION*.

5. Un profesor realizó el siguiente experimento, le preguntó a 44 alumnos que calculasen de forma aproximada, en metros, el ancho de la clase. Obtuvo las siguientes respuestas:

Grupo1								
8	9	10	10	10	10	10	10	11
11	11	11	12	12	13	13	13	14
14	14	15	15	15	15	15	15	15
15	16	16	16	17	17	17	17	18
18	20	22	25	27	35	38	40	

A otro grupo de 69 alumnos les hizo la misma pregunta pero ahora les pidió la respuesta en pies (3'28 pies = 1 metro). Ahora, las respuestas fueron:

Grupo2								
24	25	27	30	30	30	30	30	30
32	32	33	34	34	34	35	35	36
36	36	37	37	40	40	40	40	40
40	40	40	40	41	41	42	42	42
42	43	43	44	44	44	45	45	45
45	45	45	46	46	47	48	48	50
50	50	51	54	54	54	55	55	60

El ancho del aula era de 13'1 metros (43'0 pies). En base a estos datos,

1. Hacer un estudio descriptivo de estas dos muestras. Realice diagrama de cajas y bigotes y/o lo que estime conveniente para una mejor presentación de datos.
2. Calcular intervalos de confianza al 95% para la media y la varianza de la primera muestra.
3. Calcular intervalos de confianza al 95% para la diferencia de medias y el cociente de varianzas de las dos muestras.
4. ¿Puede afirmarse que el error en la aproximación es igual si se hace en metros que en pies?”