

Taller 1 – 2017

Primera serie de problemas

Para cada una de las siguientes situaciones, se pide que hagan un análisis y un informe escrito. El espíritu es simular situaciones de consultoría. A diferencia de un ejercicio teórico, aquí no se supone a priori ningún modelo; ni se recomienda ningún método. No existe “la solución verdadera” de ningún problema. Inclusive, un problema puede no tener ninguna solución enteramente satisfactoria. Enfoques distintos pueden ser aceptables. Se recomienda tener claro qué suposiciones usan, y tratar de verificar si los datos las confirman (o más bien, si no las refutan).

Consejos para la presentación

1. La comprensión, no siempre es una función monótona de la cantidad de información. Es necesario seleccionar lo esencial; no hace falta poner TODOS los números. Por ejemplo, si hay 200 datos, y se hace detección de outliers, no hace falta poner la lista de los 200 residuos, sino mostrar los valores importantes.

No mostrar más decimales que los necesarios. Decir que la pendiente es 4.58790453 con un valor-p de 0.014789345 queda poco serio. Generalmente, 3 dígitos significativos alcanzan.

2. Siempre debiera quedar claro cuál fue la selección definitiva de modelo y método; la precisión del resultado (por ejemplo, error de predicción); así como características llamativas (como outliers).

3. No hace falta poner en detalle todas las variantes que se consideraron; pero sí mencionar las que descartaron y por qué.

4. Cuando se corren varias regresiones (excluyendo distintas variables y/o observaciones, y/o transformando), que quede claro a qué corresponde cada resultado expuesto. Por ejemplo, si se corre una regresión con la “y” original, y luego con su logaritmo, cuando aparece un “R cuadrado” o un gráfico de residuos, debe saberse a cuál variante corresponde.

5. Un detalle: en los gráficos de residuos, conviene (si se puede) mostrar la línea del 0, para tener más claro el tamaño, y también para detectar asimetría y otras estructuras.

6. Identificar claramente qué variables se grafican (y cuál es la abscisa y la ordenada).

7. Conviene numerar las páginas (es más fácil hacer referencias).

8. En el nombre del archivo, identificar el remitente y el contenido (por ejemplo Juan_TP3_Ej1).

Problemas

1. Se dan las duraciones (medidas en ciclos hasta la ruptura) de una muestra de rodamientos (“rulemanes”). Describir las características principales de la muestra (posición, dispersión, asimetría), y buscar una distribución adecuada.

17.88 28.92 33.00 41.52 42.12
45.60 48.48 51.84 51.96 54.12
55.56 67.80 68.64 68.88 84.12
93.12 98.64 105.12 105.84 127.92
128.04 173.40

2. Se dan: el punto de ebullición del agua (PE) (en grados Fahrenheit) y la presión atmosférica (PA) (en pulgadas de mercurio), medidos a distintas alturas en los Alpes. Plantear un modelo que describa cómo varía PE en función de PA. ¿Con cuánta precisión se puede estimar PE en función de PA?. Comentar cualquier característica de los datos.

PE	PA
194.5	20.79
194.3	20.79
197.9	22.40
198.4	22.67
199.4	23.15
199.9	23.35
200.9	23.89
201.1	23.99
201.4	24.02
201.3	24.01
203.6	25.14
204.6	26.57
209.5	28.49
208.6	27.76
210.7	29.04
211.9	29.88
212.2	30.06

3. Se investiga el efecto de la presión aplicada durante la manufactura del papel, en el “factor de ruptura” (la fuerza necesaria para desgarrarlo). Bajo cada valor de la presión P, se manufacturó un lote de papel; de cada lote se eligieron 4 hojas, a cada una de las cuales se midió el factor de ruptura R. Se desea predecir R en función de P.

P	R			
35.0	112	119	117	113
49.5	108	99	112	118
70.0	120	106	102	109
99.0	110	101	99	104
140.0	100	102	96	101

4. La siguiente tabla da, para 12 huevos de gallina, la longitud L (o sea, el mayor diámetro), la mayor sección circular (el mayor diámetro perpendicular a L), ambas en pulgadas; y el volumen V . Interesa predecir V en función de L y M .

L	M	V
2.151	1.889	5.755
2.086	1.859	5.479
2.099	1.874	5.551
2.138	1.874	5.686
2.195	1.866	5.755
2.125	1.851	5.339
2.170	1.851	5.618
2.164	1.866	5.551
2.201	1.843	5.551
2.151	1.835	5.551
2.157	1.851	5.551
2.112	1.866	5.618